

STIC-Biotech/ChemLib

122353

From: Hutzell, Paula  
Sent: Tuesday, May 18, 2004 12:34 PM  
To: Graser, Jennifer; STIC-Biotech/ChemLib  
Subject: RE: rush search

CRF

please rush

-----Original Message-----

From: Graser, Jennifer  
Sent: Tuesday, May 18, 2004 9:37 AM  
To: STIC-Biotech/ChemLib; Hutzell, Paula  
Subject: FW: rush search  
Importance: High

Paula, this must have got lost somehow. Can you please authorize the rush search below. Thanks!

-----Original Message-----

From: STIC-Biotech/ChemLib  
Sent: Tuesday, May 18, 2004 9:35 AM  
To: Graser, Jennifer  
Subject: RE: rush search

I don't see this search in the search log..are you sure Paula approved it and sent the search to Biotech? Linda

-----Original Message-----

From: Graser, Jennifer  
Sent: Tuesday, May 18, 2004 9:03 AM  
To: STIC-Biotech/ChemLib  
Subject: FW: rush search  
Importance: High

Hi, Just checking on the status of this rush search. Thanks!

-----Original Message-----

From: Graser, Jennifer  
Sent: Tuesday, May 11, 2004 5:02 PM  
To: Hutzell, Paula  
Subject: rush search  
Importance: High

Hi Paula,

Will you please authorize a rush search for an Election which just got transferred to me?

Thanks,  
Jennifer

-----  
STIC:

**Please search amino acids 1-174 of SEQ ID NO:2; fragments from SEQ ID NO:2 which are at least 30 contiguous amino acids in length; and fragments from SEQ ID NO:2 which are at least 50 contiguous amino acids in length- from Serial No. 10/068,956.**

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:05:18 ; Search time 57 Seconds  
(without alignments)  
862.513 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 941  
Sequence: 1 RGHVVGHTLGHNSRGFV.....SAYAASAPQTQPACFPFSS 174

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	941	100.0	174	4	AA72663	Aay72663 Human pep
2	941	100.0	634	4	AAG65915	Aag65915 Amino aci
3	941	100.0	634	6	ADA54695	Ada54695 Human pro
4	483.5	51.4	363	3	AA794863	Aay94863 Human pro
5	483.5	51.4	576	4	AAG65916	Aag65916 Amino aci
6	483.5	51.4	576	5	ABB78298	Abb78298 Amino aci
7	483.5	51.4	576	6	ABR57566	Ab57566 Human Pep
8	483.5	51.4	576	7	ADE07869	Ade07869 Novel pro
9	403	42.8	530	4	AA72664	Aay72664 Murine pe
10	198.5	21.1	339	6	ABR43216	Ab43216 Human IRA
11	198.5	21.1	341	6	ABR57567	Ab57567 Human Pep
12	198.5	21.1	341	7	ADE40166	Ade40166 Human NOV
13	198.5	21.1	368	7	ADE40164	Ade40164 Human NOV
14	197.5	21.0	203	4	ABB70267	Abb70267 Drosophil
15	195.5	20.8	173	2	AAW37834	Aaw37834 Recombina
16	195.5	20.8	173	2	AAW37836	Aaw37836 Amino aci
17	195.5	20.8	196	2	AAW37837	Aaw37837 Amino aci
18	195.5	20.8	196	2	AAW37835	Aaw37835 Amino aci
19	193	20.5	191	2	AA700771	Aay00771 Human tag
20	193	20.5	196	3	AAW25583	Aab25583 Htag7 pro
21	193	20.5	196	3	AAW24022	Aab24022 Human PRO
22	193	20.5	196	3	AA799400	Aay99400 Human PRO
23	193	20.5	196	3	AA796964	Aay96964 Chondrosa
24	193	20.5	196	4	AAW66149	Aab66149 Protein o
25	193	20.5	196	6	ABO33642	Abo33642 Novel hum

26	193	20.5	196	6	ADA27055	Ada27055 Human nov
27	193	20.5	196	7	ABO44495	Abo44495 Human sec
28	193	20.5	196	7	ABR62398	Abr62398 Polypepti
29	193	20.5	196	7	ABO33519	Abo33519 Novel hum
30	193	20.5	196	7	ADC18085	Adc18085 Human PRO
31	193	20.5	196	7	ADD70731	Add70731 Human sec
32	193	20.5	196	7	ADD39808	Add39808 Human sec
33	193	20.5	196	7	ADD70254	Add70254 Human sec
34	193	20.5	196	7	ADD38375	Add38375 Human sec
35	193	20.5	196	7	ADD39331	Add39331 Human sec
36	193	20.5	196	7	ADD38854	Add38854 Human sec
37	193	20.5	196	7	ADD40285	Add40285 Human sec
38	193	20.5	196	7	ADE50506	Ade50506 Human sec
39	193	20.5	196	7	ADE20118	Ade20118 Human sec
40	193	20.5	196	7	ADE50029	Ade50029 Human sec
41	193	20.5	196	7	ADE21587	Ade21587 Human sec
42	193	20.5	196	8	ADE86585	Ade86585 Novel hum
43	188.5	20.0	369	6	ABJ19367	Abj19367 NOVX rela
44	183.5	19.5	241	5	ABB53271	Abb53271 Human pol
45	183.5	19.5	368	3	AAY96963	Aay96963 Wound hea

ALIGNMENTS

RESULT 1  
AA72663  
ID AA72663 standard; protein; 174 AA.

XX AC AA72663;

XX 31-MAY-2001 (first entry)

Human peptidoglycan recognition protein-related liver protein.

Peptidoglycan recognition protein-related liver protein; PGRP-L; food additive; food preservative; breast cancer; ovarian cancer; immune disorder; Addison's disease; allergy; cardiovascular disorder; myocardial ischaemia; wound healing; neurological disease; vasotropic; Alzheimer's disease; immunosuppressive; antiarthritic; antineumatic; antiproliferative; cytostatic; cerebroprotective; antibacterial; virucide; fungicide; ophthalmological; human; gene therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT	Region	12-13	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	34..47	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	51..57	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	64..75	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	80..84	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	96..133	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	135..137	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	145..154	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	161..167	/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"



Region 172. .174  
/note= "Antigenic region of human peptidoglycan  
recognition protein-related liver protein"

WO200114545-A1.  
01-MAR-2001.  
18-AUG-2000; 2000WO-US022877.  
20-AUG-1999; 99US-0149715P.  
(HUMA-) HUMAN GENOME SCI INC.  
Young PE, Rosen CA, Duan RD;  
WPI; 2001-160115/16.  
N-PSDB; AAD02742.

Isolated nucleic acids encoding human and murine peptidoglycan  
recognition protein-related liver (PGRP-L) proteins, useful for  
preventing or treating autoimmune diseases e.g. rheumatoid arthritis, and  
hyperproliferative disorders.

Claim 11; Fig 1; 291pp; English.

The invention relates to human and murine peptidoglycan recognition  
protein-related liver proteins (PGRP-L) and nucleic acid molecules  
encoding them. The polypeptides of the invention can be used as food  
additive or preservative to increase or decrease storage capabilities.  
The PGRP-L polynucleotides are used for chromosome identification. They  
are also useful as probes for diagnosing disorders related to the female  
reproductive system, particularly breast and ovary cancer. They are also  
useful in the gene therapy of breast and ovarian cancer. The PGRP-L  
polynucleotides, polypeptides, and their antibodies, agonists and  
antagonists are useful in the diagnosis, treatment and prevention of  
cancer particularly breast and ovarian cancer, and cancers of the adrenal  
gland, gastrointestinal tract, liver, lung, or urogenital; immune  
disorders such as Addison's disease, allergies, autoimmune haemolytic  
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
cardiovascular disorders such as myocardial ischaemias; wound healing;  
neurological diseases such as Alzheimer's disease, cerebral anoxia and  
epilepsy; and infectious diseases such as viral, bacterial, fungal and  
parasitic infections. The present sequence is human peptidoglycan  
recognition protein- related liver protein (hPGRP-L). This protein has  
molecular weight of about 18,595 Da

Sequence 174 AA;  
Query Match 100.0%; Score 941; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 2.7e-88;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGHWVGAHTLGHNSRGFGVAIVGNVTAALPTEAALRTVTRDTLPSCAVRAGLLRPDYALL 60  
Db 1 RGHWVGAHTLGHNSRGFGVAIVGNVTAALPTEAALRTVTRDTLPSCAVRAGLLRPDYALL 60  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPVSVTSTRPLPPACNSCART 120  
Db 61 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPVSVTSTRPLPPACNSCART 120  
QY 121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQOTQACFPSS 174  
Db 121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQOTQACFPSS 174

RESULT 2  
AAG65915  
ID AAG65915 standard; protein; 634 AA.  
XX  
AC AAG65915;  
XX

DT 11-FEB-2002 (first entry)  
XX  
DE Amino acid sequence of GSK gene Id 239881.  
XX  
KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;  
KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;  
KW cyostatic; cerebroprotective; vasotropic; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200172961-A2.  
PD 04-OCT-2001.  
XX  
PF 22-MAR-2001; 2001WO-US009226.  
XX  
PR 24-MAR-2000; 2000US-0192158P.  
PR 28-MAR-2000; 2000US-0192668P.  
PR 27-APR-2000; 2000US-0200166P.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y;  
XX  
DR WPI; 2001-639223/73.  
DR N-PSDB; AAI67205.  
XX  
PT Isolated polypeptides, which may be peptide hormones, which are  
PT identified by high throughput genome-based biology which identifies genes  
PT and gene products as therapeutic targets for treatment of diseases such  
PT as diabetes and cancer.  
XX  
PS Claim 1; Page 94-95; 99pp; English.

The invention provides polypeptides (AAG65886-65918) which may be peptide  
hormones (including insulin, growth hormones, chemokines, cytokines,  
neuropeptides, integrins, kallikreins, lamins, melanins, natruiretic  
hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
secretogranins, selectins, thromboglobulins, thymosins) identified by  
high throughput genome-based biology and polynucleotides (AAI67176-67208)  
encoding them. The polypeptides can be expressed by standard recombinant  
methodology. The polypeptides are useful in the treatment of disease such  
as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
asthma, manic depression, dementia, delirium, mental retardation,  
Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental  
or sexual development disorders, and dysfunctions of the blood cascade  
system including those leading to stroke. The polynucleotides may be used  
as diagnostic reagents through detecting mutations in the associated gene  
and for chromosome localization and for tissue expression studies. The  
polypeptides and polynucleotides may also be used as vaccines

Sequence 634 AA;  
Query Match 100.0%; Score 941; DB 4; Length 634;  
Best Local Similarity 100.0%; Pred. No. 1.3e-87;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGHWVGAHTLGHNSRGFGVAIVGNVTAALPTEAALRTVTRDTLPSCAVRAGLLRPDYALL 60  
Db 461 RGHWVGAHTLGHNSRGFGVAIVGNVTAALPTEAALRTVTRDTLPSCAVRAGLLRPDYALL 520  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPVSVTSTRPLPPACNSCART 120  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPVSVTSTRPLPPACNSCART 580  
QY 121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQOTQACFPSS 174  
Db 581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQOTQACFPSS 634

*Offen*

RESULT 3  
ADA54695  
ID ADA54695 standard; protein; 634 AA.  
XX  
AC ADA54695;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human protein, SEQ ID 2263.  
XX  
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
KW inflammatory disease; osteoporosis; neurological disease.  
XX  
OS Homo sapiens.  
XX  
PN ~~EP1293569-A2~~  
XX  
PD 19-MAR-2003.  
XX  
PF 21-MAR-2002; 2002EP-00006586.  
XX  
PR 14-SEP-2001; 2001JP-00328381.  
PR 24-JAN-2002; 2002US-0350435P.  
XX  
PA (HELI-) HELIX RES INST.  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
DR WPI; 2003-395539/38.  
DR N-PSDB; ADA53056.  
XX  
PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.  
XX  
PS Claim 14; SEQ ID NO 2263; 205pp; English.  
XX  
CC The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.  
XX  
SQ Sequence 634 AA;  
Query Match 100.0%; Score 941; DB 6; Length 634;  
Best Local Similarity 100.0%; Pred. No. 1.3e-87;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGHWVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60  
Db 461 RGHWVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 120  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 580  
QY 121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYASAQPQTQPCFPSS 174  
Db 581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYASAQPQTQPCFPSS 634  
RESULT 4  
AAY94863  
ID AAY94863 standard; protein; 363 AA.  
XX  
AC AAY94863;  
XX  
DT 12-JUN-2000 (first entry)

XX Human protein clone HP10477.  
DE  
XX Human protein; hydrophobic domain; nutritional source; haematopoiesis;  
KW cytokine production; cell proliferation; cell differentiation;  
KW immune deficiency; infectious disease; autoimmune disorder; asthma;  
KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;  
KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;  
KW nervous system disorder; Alzheimer's disease; Parkinson's disease;  
KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;  
KW systemic cytokine damage; tissue differentiation; contraceptive; stroke;  
KW coagulation disorder; myocardial infarction; inflammatory condition;  
KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;  
KW nephritis; therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200005367-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 22-JUL-1999; 99WO-JP003929.  
XX  
PR 24-JUL-1998; 98JP-00208820.  
PR 07-AUG-1998; 98JP-00224105.  
PR 25-AUG-1998; 98JP-00238116.  
PR 09-SEP-1998; 98JP-00254736.  
PR 29-SEP-1998; 98JP-00275505.  
XX  
PA (SAGA ) SAGAMI CHEM RES CENT.  
PA (PROT-) PROTEGENE INC.  
PI Kato S, Kimura T;  
XX  
DR WPI; 2000-182694/16.  
XX  
PT Novel human proteins having hydrophobic domains useful for treating  
PT osteoporosis, Alzheimer's disease, Parkinson's disease, asthma, multiple  
PT sclerosis, rheumatoid arthritis, cancer, anemia, and stroke.  
XX  
PS Claim 1; Page 211-212; 351pp; English.  
XX  
CC This sequence represents a human protein of the invention, which has  
CC hydrophobic domains. The DNA sequences can be used as a probe or as a  
CC genetic marker. The protein can also be used as a marker, and to identify  
CC potential genetic disorders. The DNA and protein can also be used as  
CC nutritional sources or supplements. The protein exhibits cytokine, cell  
CC proliferation, cell differentiation activities and induces production of  
CC other cytokines in certain cell populations. The protein also exhibits  
CC immune stimulating or immune suppressing activity. It can be used in the  
CC treatment of various immune deficiencies and disorders, and to treat  
CC infectious diseases caused by viral, bacterial, fungal or other  
CC infections. The protein is also used for treating autoimmune disorders  
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid  
CC arthritis. It is also useful in the treatment of allergic reactions and  
CC conditions such as asthma, and in immune suppression after organ  
CC transplantation. The protein is useful in regulation of haematopoiesis  
CC and consequently in the treatment of myeloid or lymphoid cell  
CC deficiencies. It is also used in compositions for tissue growth or  
CC regeneration. The protein is also used in the treatment of osteoporosis  
CC or osteoarthritis and in the treatment of periodontal disease and other  
CC tooth repair processes. The protein is used in the treatment of nervous  
CC system disorders such as Alzheimer's disease, Parkinson's disease, and  
CC Huntington's disease. They are useful for protection or regeneration and  
CC treatment of lung or liver fibrosis, reperfusion injury in various  
CC tissues, and conditions resulting from systemic cytokine damage. They are  
CC also used for promoting or inhibiting tissue differentiation. They are  
CC also used as contraceptives since they exhibit activin or inhibin related  
CC activities and as a fertility inducing therapeutic. They are used for  
CC treating various coagulation disorders and in treatment and prevention of  
CC conditions resulting from coagulation activities e.g. myocardial  
CC infarction or stroke. They also acts as receptors, receptor ligands or  
CC inhibitors or agonists of receptor/ligand interactions. They are used to



CC treat inflammatory conditions such as septic shock, sepsis, ischaemia  
CC reperfusion injury, arthritis, and nephritis. They can be used to prevent  
CC tumours  
XX  
SQ Sequence 363 AA;  
Query Match 51.4%; Score 483.5; DB 3; Length 363;  
Best Local Similarity 83.3%; Pred. No. 5.7e-41;  
Matches 95; Conservative 0; Mismatches 10; Indels 9; Gaps 2;  
QY 1 RGHVWGHTLGHNSRGFGVAIVGNVTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60  
Db 248 RGHVWGHTLGHNSRGFGVAIVGNVTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 307  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRP--SVYTSSTRPLPP 112  
Db 308 GHRQLVRTDCPGDALFDLLRTWPHTAT-----VKPRPARSVSKRSRREPPP 354  
RESULT 5  
AAG65916  
ID AAG65916 standard; protein; 576 AA.  
XX  
AC AAG65916;  
XX  
DT 11-FEB-2002 (first entry)  
XX  
DE Amino acid sequence of GSK gene Id 239881.  
XX  
KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;  
KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;  
KW cyostatic; cerebroprotective; vasotropic; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200172961-A2  
XX  
PD 04-OCT-2001.  
XX  
PF 22-MAR-2001; 2001WO-US009226.  
XX  
PR 24-MAR-2000; 2000US-0192158P.  
PR 28-MAR-2000; 2000US-0192668P.  
PR 27-APR-2000; 2000US-0200166P.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y;  
XX  
DR WPI; 2001-639223/73.  
DR N-PSDB; AAI67206.  
XX  
PT Isolated polypeptides, which may be peptide hormones, which are  
PT identified by high throughput genome-based biology which identifies genes  
PT and gene products as therapeutic targets for treatment of diseases such  
PT as diabetes and cancer.  
XX  
PS Claim 1; Page 95-96; 99pp; English.  
XX  
CC The invention provides polypeptides (AAG65886-65918) which may be peptide  
CC hormones (including insulin, growth hormones, chemokines, cytokines,  
CC neuropeptides, integrins, kallikreins, lamins, melanins, natriuretic  
CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
CC secretogranins, selectins, thromboglobulins, thymosins) identified by  
CC high throughput genome-based biology and polynucleotides (AAI67176-67208)  
CC encoding them. The polypeptides can be expressed by standard recombinant  
CC methodology. The polypeptides are useful in the treatment of disease such  
CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
CC asthma, manic depression, dementia, delirium, mental retardation,  
CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental

CC or sexual development disorders, and dysfunctions of the blood cascade  
CC system including those leading to stroke. The polynucleotides may be used  
CC as diagnostic reagents through detecting mutations in the associated gene  
CC and for chromosome localization and for tissue expression studies. The  
CC polypeptides and polynucleotides may also be used as vaccines  
XX  
SQ Sequence 576 AA;  
Query Match 51.4%; Score 483.5; DB 4; Length 576;  
Best Local Similarity 83.3%; Pred. No. 1e-40;  
Matches 95; Conservative 0; Mismatches 10; Indels 9; Gaps 2;  
QY 1 RGHVWGHTLGHNSRGFGVAIVGNVTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60  
Db 461 RGHVWGHTLGHNSRGFGVAIVGNVTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRP--SVYTSSTRPLPP 112  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHTAT-----VKPRPARSVSKRSRREPPP 567  
RESULT 6  
ABB78298  
ID ABB78298 standard; protein; 576 AA.  
XX  
AC ABB78298;  
XX  
DT 05-DEC-2002 (first entry)  
XX  
DE Amino acid sequence of lipid-associated molecule (LIPAM)-9.  
XX  
KW Human; lipid-associated molecule; LIPAM; cardiovascular disease;  
KW atherosclerosis; hypertension; aneurysm; congestive heart failure;  
KW angina pectoris; heart disease; lung disease; oedema; emphysema;  
KW bronchitis; gastrointestinal disease; nausea; peptic ulcer;  
KW Crohn's disease; lipid metabolism; Fabry's disease; diabetes mellitus;  
KW hyperlipidaemia; autoimmune disease; inflammatory disease;  
KW acquired immunodeficiency syndrome; AIDS; anaemia; asthma; gout;  
KW pancreatitis; neurological disease; stroke; Alzheimer's disease;  
KW multiple sclerosis; Parkinson's disease; anxiety; schizophrenia; amnesia;  
KW metabolic disease; Addison's disease; developmental disease;  
KW Cushing's syndrome; endocrine disease; cell proliferative disorder;  
KW cancer; leukemia; lymphoma; sarcoma.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Modified-site 77 Location/Qualifiers  
FT /note= "potential glycosylation site"  
FT Modified-site 79  
FT /note= "potential phosphorylation site"  
FT Modified-site 154  
FT /note= "potential phosphorylation site"  
FT Modified-site 181  
FT /note= "potential phosphorylation site"  
FT Modified-site 202  
FT /note= "potential phosphorylation site"  
FT Modified-site 213  
FT /note= "potential phosphorylation site"  
FT Domain 215. .234  
FT /note= "transmembrane domain"  
FT Modified-site 239  
FT /note= "potential phosphorylation site"  
FT Domain 255. .283  
FT /note= "transmembrane domain"  
FT Modified-site 259  
FT /note= "potential phosphorylation site"  
FT Modified-site 367  
FT /note= "potential glycosylation site"  
FT Modified-site 485  
FT /note= "potential glycosylation site"  
FT Modified-site 498  
FT /note= "potential phosphorylation site"



FT Modified-site 548 /note= "potential phosphorylation site"  
FT Modified-site 558 /note= "potential phosphorylation site"  
FT Modified-site 561 /note= "potential phosphorylation site"  
XX

PN WO200263005-A2.

XX  
PD 15-AUG-2002.

XX  
PF 06-FEB-2002; 2002WO-US003813.

XX  
PR 06-FEB-2001; 2001US-0266910P.

PR 16-MAR-2001; 2001US-0276855P.

PR 16-MAR-2001; 2001US-0276891P.

PR 28-MAR-2001; 2001US-0279760P.

PR 13-APR-2001; 2001US-0283818P.

PR 20-APR-2001; 2001US-0285405P.

XX  
PA (INCY-) INCYTE GENOMICS INC.

XX  
PI Das D, Yao MG, Arvizu C, Baughn MR, Lu Y, Hafalia AJA, Walia NK;

PI Griffin JA, Lu DAM, Yue H, Ding L, Townley DJ, Elliott VS;

PI Forsythe I, Ramkumar J, Gandhi AR, Ison CH, Warren BA, Tang YT;

PI Emerling BM, Honchell CD;

XX  
DR WPI; 2002-627558/67.

DR N-PSDB; ABV72135.

XX  
XX  
PS  
XX

XX  
PT New human lipid-associated molecules, useful for diagnosing, treating or

PT preventing cardiovascular disorders (e.g. aneurysms), neurological

PT disorders (e.g. Parkinson's disease) or cancers (e.g. leukemia or

PT lymphoma).

XX  
XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

FT Modified-site 548 /note= "potential phosphorylation site"  
FT Modified-site 558 /note= "potential phosphorylation site"  
FT Modified-site 561 /note= "potential phosphorylation site"  
XX

PN WO200263005-A2.

XX  
PD 15-AUG-2002.

XX  
PF 06-FEB-2002; 2002WO-US003813.

XX  
PR 06-FEB-2001; 2001US-0266910P.

PR 16-MAR-2001; 2001US-0276855P.

PR 16-MAR-2001; 2001US-0276891P.

PR 28-MAR-2001; 2001US-0279760P.

PR 13-APR-2001; 2001US-0283818P.

PR 20-APR-2001; 2001US-0285405P.

XX  
PA (INCY-) INCYTE GENOMICS INC.

XX  
PI Das D, Yao MG, Arvizu C, Baughn MR, Lu Y, Hafalia AJA, Walia NK;

PI Griffin JA, Lu DAM, Yue H, Ding L, Townley DJ, Elliott VS;

PI Forsythe I, Ramkumar J, Gandhi AR, Ison CH, Warren BA, Tang YT;

PI Emerling BM, Honchell CD;

XX  
DR WPI; 2002-627558/67.

DR N-PSDB; ABV72135.

XX  
XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

FT Modified-site 548 /note= "potential phosphorylation site"  
FT Modified-site 558 /note= "potential phosphorylation site"  
FT Modified-site 561 /note= "potential phosphorylation site"  
XX

PN WO200263005-A2.

XX  
PD 15-AUG-2002.

XX  
PF 06-FEB-2002; 2002WO-US003813.

XX  
PR 06-FEB-2001; 2001US-0266910P.

PR 16-MAR-2001; 2001US-0276855P.

PR 16-MAR-2001; 2001US-0276891P.

PR 28-MAR-2001; 2001US-0279760P.

PR 13-APR-2001; 2001US-0283818P.

PR 20-APR-2001; 2001US-0285405P.

XX  
PA (INCY-) INCYTE GENOMICS INC.

XX  
PI Das D, Yao MG, Arvizu C, Baughn MR, Lu Y, Hafalia AJA, Walia NK;

PI Griffin JA, Lu DAM, Yue H, Ding L, Townley DJ, Elliott VS;

PI Forsythe I, Ramkumar J, Gandhi AR, Ison CH, Warren BA, Tang YT;

PI Emerling BM, Honchell CD;

XX  
DR WPI; 2002-627558/67.

DR N-PSDB; ABV72135.

XX  
XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

FT Modified-site 548 /note= "potential phosphorylation site"  
FT Modified-site 558 /note= "potential phosphorylation site"  
FT Modified-site 561 /note= "potential phosphorylation site"  
XX

PN WO200263005-A2.

XX  
PD 15-AUG-2002.

XX  
PF 06-FEB-2002; 2002WO-US003813.

XX  
PR 06-FEB-2001; 2001US-0266910P.

PR 16-MAR-2001; 2001US-0276855P.

PR 16-MAR-2001; 2001US-0276891P.

PR 28-MAR-2001; 2001US-0279760P.

PR 13-APR-2001; 2001US-0283818P.

PR 20-APR-2001; 2001US-0285405P.

XX  
PA (INCY-) INCYTE GENOMICS INC.

XX  
PI Das D, Yao MG, Arvizu C, Baughn MR, Lu Y, Hafalia AJA, Walia NK;

PI Griffin JA, Lu DAM, Yue H, Ding L, Townley DJ, Elliott VS;

PI Forsythe I, Ramkumar J, Gandhi AR, Ison CH, Warren BA, Tang YT;

PI Emerling BM, Honchell CD;

XX  
DR WPI; 2002-627558/67.

DR N-PSDB; ABV72135.

XX  
XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

FT Modified-site 548 /note= "potential phosphorylation site"  
FT Modified-site 558 /note= "potential phosphorylation site"  
FT Modified-site 561 /note= "potential phosphorylation site"  
XX

PN WO200263005-A2.

XX  
PD 15-AUG-2002.

XX  
PF 06-FEB-2002; 2002WO-US003813.

XX  
PR 06-FEB-2001; 2001US-0266910P.

PR 16-MAR-2001; 2001US-0276855P.

PR 16-MAR-2001; 2001US-0276891P.

PR 28-MAR-2001; 2001US-0279760P.

PR 13-APR-2001; 2001US-0283818P.

PR 20-APR-2001; 2001US-0285405P.

XX  
PA (INCY-) INCYTE GENOMICS INC.

XX  
PI Das D, Yao MG, Arvizu C, Baughn MR, Lu Y, Hafalia AJA, Walia NK;

PI Griffin JA, Lu DAM, Yue H, Ding L, Townley DJ, Elliott VS;

PI Forsythe I, Ramkumar J, Gandhi AR, Ison CH, Warren BA, Tang YT;

PI Emerling BM, Honchell CD;

XX  
DR WPI; 2002-627558/67.

DR N-PSDB; ABV72135.

XX  
XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

XX  
PS  
XX

QY 61 GHRQLVRTDCPGDALFDLLRTWPHFTAVSLRSLHYTARRP--SVYTSSTRPLPP 112  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHFTAT-----VKRPARSVSKSRREPPP 567

RESULT 8  
ADE07869  
ID ADE07869 standard; protein; 576 AA.  
XX  
AC ADE07869;  
XX

DT 29-JAN-2004 (first entry)  
XX  
DE Novel protein (useful for identifying genetic disorders) #24.  
XX

XX novel gene; novel protein; tissue marker; molecular weight marker;  
KW chromosome marker; genetic disorder.  
XX

OS Unidentified.  
XX

XX W02003054152-22  
PD 03-JUL-2003.  
XX

PF 10-DEC-2002; 2002WO-US039555.  
XX

PR 10-DEC-2001; 2001US-0339739P.  
PR 11-DEC-2001; 2001US-0339453P.  
PR 14-MAR-2002; 2002US-0365091P.  
PR 14-MAR-2002; 2002US-0365384P.  
PR 12-APR-2002; 2002US-0372381P.  
PR 12-APR-2002; 2002US-0372615P.  
PR 22-APR-2002; 2002US-00128558.  
PR 24-APR-2002; 2002US-0376045P.  
XX

PA (HYSE-) HYSEQ INC.  
XX

PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
XX

DR WPI; 2003-569235/53.  
DR N-PSDB; ADE06958.  
XX

PT New polynucleotides, useful for expressing recombinant proteins for  
PT analysis, characterization or therapeutic use, or as markers for tissues  
PT in which the corresponding protein is preferentially expressed.  
XX

PS Claim 20; SEQ ID NO 935; 1177pp; English.  
XX

CC The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present amino acid sequence represents a protein  
CC of the invention.  
XX

SQ Sequence 576 AA;

Query Match 51.4%; Score 483.5; DB 7; Length 576;  
Best Local Similarity 83.3%; Pred. No. 1e-40;  
Matches 95; Conservative 0; Mismatches 10; Indels 9; Gaps 2;

QY 1 RGHWVGAHTLGHNSRGFVAIGNYTAALPTEAALRTVRDITLPSCAVRAGLLRPDYALL 60  
Db 461 RGHWVGAHTLGHNSRGFVAIGNYTAALPTEAALRTVRDITLPSCAVRAGLLRPDYALL 520

QY 61 GHRQLVRTDCPGDALFDLLRTWPHFTAVSLRSLHYTARRP--SVYTSSTRPLPP 112  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHFTAT-----VKRPARSVSKSRREPPP 567

RESULT 9  
AA72664  
ID AA72664 standard; protein; 530 AA.  
XX  
AC AA72664;  
XX  
DT 31-MAY-2001 (first entry)  
XX  
DE Murine peptidoglycan recognition protein-related liver protein.  
XX  
KW Peptidoglycan recognition protein-related liver protein; PGRP-L;  
KW food additive; food preservative; breast cancer; ovarian cancer;  
KW immune disorder; Addison's disease; allergy; cardiovascular disorder;  
KW myocardial ischaemia; wound healing; neurological disease; vasotropic;  
KW Alzheimer's disease; immunosuppressive; antiarthritic; antineumatic;  
KW antiproliferative; cytostatic; cerebroprotective; antibacterial;  
KW virucide; fungicide; ophthalmological; murine; gene therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1. .22  
FT /label= Signal\_peptide  
FT Region 18. .24  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Protein 23. .530  
FT /label= Mature\_murine\_peptidoglycan\_recognition\_protein-  
FT related\_liver\_protein  
FT Region 37. .45  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 59. .68  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 76. .93  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 101. .111  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 116. .122  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 143. .159  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 172. .197  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 213. .229  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 242. .246  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 268. .275  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 288. .305  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 325. .330  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 335. .340  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"  
FT Region 348. .353  
FT /note= "Antigenic region of murine peptidoglycan  
FT recognition protein-related liver protein"



Region 362. .382  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
397. .399  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
406. .413  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
415. .426  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
432. .435  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
440. .442  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
449. .457  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
466. .475  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
478. .484  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
486. .499  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
511. .513  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
521. .530  
/note= "Antigenic region of murine peptidoglycan  
recognition protein-related liver protein"  
WO200114545-A1.  
01-MAR-2001.  
18-AUG-2000; 2000WO-US022877.  
20-AUG-1999; 99US-0149715P.  
(HUMA-) HUMAN GENOME SCI INC.  
Young PE, Rosen CA, Duan RD;  
WPI; 2001-160115/16.  
N-PSDB; AAD02743.  
Isolated nucleic acids encoding human and murine peptidoglycan  
recognition protein-related liver (PGRP-L) proteins, useful for  
preventing or treating autoimmune diseases e.g. rheumatoid arthritis, and  
hyperproliferative disorders.  
Disclosure; Fig 2; 291pp; English.  
The invention relates to human and murine peptidoglycan recognition  
protein-related liver proteins (PGRP-L) and nucleic acid molecules  
encoding them. The polypeptides of the invention can be used as food  
additive or preservative to increase or decrease storage capabilities.  
The PGRP-L polynucleotides are used for chromosome identification. They  
are also useful as probes for diagnosing disorders related to the female  
reproductive system, particularly breast and ovary cancer. They are also  
useful in the gene therapy of breast and ovarian cancer. The PGRP-L  
polynucleotides, polypeptides, and their antibodies, agonists and  
antagonists are useful in the diagnosis, treatment and prevention of  
cancer particularly breast and ovarian cancer, and cancers of the adrenal  
gland, gastrointestinal tract, liver, lung, or urogenital; immune  
disorders such as Addison's disease, allergies, autoimmune haemolytic  
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischaemias; wound healing;  
CC neurological diseases such as Alzheimer's disease, cerebral anoxia and  
CC epilepsy; and infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections. The present sequence is murine peptidoglycan  
CC recognition protein- related liver protein (mPGRP-L). This protein has  
CC molecular weight of about 57,764 Da  
XX  
SQ Sequence 530 AA;  
Query Match 42.8%; Score 403; DB 4; Length 530;  
Best Local Similarity 81.8%; Pred. No. 1.6e-32;  
Matches 72; Conservative 5; Mismatches 11; Indels 0; Gaps 0;  
Qy 1 RGHVWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVDTLPSCAVRAGLLRPDYALL 60  
Db 441 RGHVWVGAHTRGYNSRGFGVAFVGNVTGSLPNEAALNTVRDALPSCAIREGLLRPDYKLL 500  
Qy 61 GHRQLVVRTDCPGDALFDLLRTWPHFTAV 88  
Db 501 GHRQLVLVTHCPGNALFNLLRTWPHFTTEV 528  
RESULT 10  
ABR43216  
ID ABR43216 standard; protein; 339 AA.  
XX  
AC ABR43216;  
XX  
DT 07-JUL-2003 (first entry)  
XX  
DE Human IRAP-12 protein SEQ ID NO:12.  
XX  
KW Human; immune response associated protein; IRAP; anti-HIV; antiallergic;  
KW antiinflammatory; antianaemic; antiparkinsonian; nootropic; antithyroid;  
KW anticonvulsant; antiarteriosclerotic; antiasthmatic; immunosuppressive;  
KW cytostatic; hepatotropic; dermatological; antidiabetic; nephrotropic;  
KW antigout; thyromimetic; neuroprotective; osteopathic; antiarthritic;  
KW antiparasitic; antihelminthic; antipsoriatic; uropathic; protozoacide;  
KW antirheumatic; haemostatic; antibacterial; virucide; ophthalmological;  
KW fungicide; gene therapy; immune system disorder; neurological disorder;  
KW developmental disorder; muscle disorder; cell proliferative disorder;  
XX  
OS Homo sapiens.  
XX  
PW WO2003025542-A2.  
PD 27-MAR-2003.  
XX  
PF 19-SEP-2002; 2002WO-US029979.  
XX  
PR 21-SEP-2001; 2001US-0324034P.  
PR 05-OCT-2001; 2001US-0327395P.  
PR 12-OCT-2001; 2001US-0328923P.  
PR 19-OCT-2001; 2001US-0342810P.  
PR 09-NOV-2001; 2001US-0344468P.  
PR 21-NOV-2001; 2001US-0332140P.  
PR 07-DEC-2001; 2001US-0340282P.  
PR 09-JAN-2002; 2002US-0347693P.  
PR 20-FEB-2002; 2002US-0358279P.  
PR 01-MAR-2002; 2002US-0361088P.  
PR 15-MAR-2002; 2002US-0364494P.  
PR 10-MAY-2002; 2002US-0379876P.  
PR 11-JUN-2002; 2002US-0388180P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Ho A, Baughn MR, Becha SD, Burford N, Elliott VS, Emerling BM;  
PI Forsythe IJ, Gorvad AE, Griffin JA, Hafalia AJA, Honchell CD;  
PI Ison CH, Burrill JD, Blake JJ, Lal PG, Lee EA, Marquis JP;  
PI Lehr-Mason PM, Lee S, Sprague WW, Swarnakar A, Tang YT, Tran B;  
PI Tran UK, Bhatia U, Chawla NK, Warren BA, Zheng W, Xu Y, Yue H;



XX WPI; 2003-363161/34.  
DR N-PSDB; ACC59935.  
XX  
PT New human immune response associated proteins and polynucleotides, useful  
PT for diagnosing, treating or preventing immune system disorders, e.g. AIDS  
PT or anemia, cell proliferative disorders, e.g. cancer, or neurological  
PT disorders.  
XX  
PS Claim 1; Page 176-177; 213pp; English.  
XX  
CC ACC59924 to ACC59958 encode the human immune response associated proteins  
CC given in ABR43205 to ABR43239, designated IRAP-1 to IRAP-29 (I). (I) have  
CC anti-HIV, anti-allergic, anti-inflammatory, antianaemic, antiparkinsonian,  
CC nootropic, anticonvulsant, antiarteriosclerotic, antiasthmatic, antigout,  
CC immunosuppressive, antithyroid, cytostatic, hepatotropic, dermatological,  
CC antidiabetic, nephrotropic, thyromimetic, neuroprotective, osteopathic,  
CC antiarthritic, antiparasitic, antihelminthic, antipsoriatic, uropathic,  
CC ophthalmological, antirheumatic, haemostatic, antibacterial, virucide,  
CC protozoacide and fungicide activities, and can be used in gene therapy.  
CC Human IRAP polynucleotides, agonists and antagonists are useful for  
CC diagnosing, treating or preventing disorders associated with aberrant  
CC expression of IRAP, such as immune system disorders (e.g. AIDS, asthma,  
CC allergies, autoimmune thyroiditis, contact dermatitis, Crohn's disease,  
CC diabetes mellitus, glomerulonephritis, Good pasture's syndrome, gout,  
CC Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,  
CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's  
CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis),  
CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or  
CC epilepsy), developmental disorders (e.g. renal tubular acidosis, anaemia  
CC or mental retardation), muscle disorder (e.g. cardiomyopathy, myocarditis  
CC or dermatolysisitis), cell proliferative disorders (e.g. arteriosclerosis,  
CC atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal  
CC haemoglobinuria, polycythaemia vera, psoriasis, primary  
CC thrombocytopenia or cancer), or viral, bacterial, fungal, parasitic,  
CC protozoan or helminthic infections  
XX  
SQ Sequence 339 AA;

Query Match 21.1%; Score 198.5; DB 6; Length 339;  
Best Local Similarity 44.0%; Pred. No. 9e-12;  
Matches 37; Conservative 12; Mismatches 34; Indels 1; Gaps 1;  
QY 2 GWHVVGHTLGHNSRGFGVAIVGNYYTAAALPTEAALRTVDRDTPSCAVRAGLLRPDYALLG 61  
Db 255 GWHIQGSHTYGFNDIALGIAFIGYFVEKPPNAAALEAAQD-LIQCAVVEGYLTPNYLLMG 313  
QY 62 HRQLVRTDCPGDALFDLLRTWPHF 85  
Db 314 HSDVNVILSPGQALYNIISTWPHF 337

RESULT 11  
ABR57567  
ID ABR57567 standard; protein; 341 AA.  
XX  
AC ABR57567;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human Peptidoglycan Recognition Protein, PGRP-Ialpha.  
XX  
KW Immunomodulator; gene therapy; human; antimicrobial; chromosome 1q;  
KW Peptidoglycan Recognition Protein; PGRP; PGRP-Long; PGRP-L;  
KW PGRP-intermediate alpha; PGRP-intermediate beta; PGRP-Ialpha; PGRP-Ibeta.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..17  
FT /label= signal\_peptide  
FT Protein 18..341  
FT /label= Mature\_protein

Domain 81..108  
FT /label= PGRP\_Domain\_IV  
FT 125..145  
FT /label= Transmembrane\_domain\_#1  
FT 197..211  
FT /label= PGRP\_Domain\_III  
FT 237..265  
FT /label= PGRP\_domain\_II  
FT 264..282  
FT /label= Transmembrane\_domain\_#2  
FT 290..339  
FT /label= PGRP\_domain\_I  
XX WO2003029401-A2.  
PN 10-APR-2003.  
XX  
PD 15-JUL-2002; 2002WO-US022428.  
XX  
PF 13-JUL-2001; 2001US-0305049P.  
PR (ADRE-) ADVANCED RES & TECHNOLOGY INST.  
XX  
PA Dziarski R, Liu C, Xu Z, Gupta D;  
PI WPI; 2003-381614/36.  
XX N-PSDB; ACC70430.  
DR  
DR New peptidoglycan recognition proteins (PGRP)-L, PGRP-I alpha and PGRP-I  
PT beta, useful for modulating the innate immune system, and as research  
PT tools for identifying other proteins involved in regulating antimicrobial  
PT processes.  
XX  
PS Claim 16; Fig 9; 86pp; English.  
XX  
CC The present invention relates to novel human Peptidoglycan Recognition  
CC Proteins (PGRP): PGRP-Long (PGRP-L), PGRP-intermediate alpha and PGRP-  
CC intermediate beta (PGRP-Ialpa and PGRP-Ibeta, ACC70429-ACC70431 and  
CC ABR57566-ABR57568). The PGRPs and their coding sequences are useful for  
CC modulating innate immune system, and as research tools to identify other  
CC proteins that are intimately involved in the regulation of antimicrobial  
CC processes. The gene for PGRP-L is located on chromosome 19 and the genes  
CC for PGRP-Ialpa and PGRP-Ibeta are located on chromosome 1q  
XX  
SQ Sequence 341 AA;  
Query Match 21.1%; Score 198.5; DB 6; Length 341;  
Best Local Similarity 44.0%; Pred. No. 9.1e-12;  
Matches 37; Conservative 12; Mismatches 34; Indels 1; Gaps 1;  
QY 2 GWHVVGHTLGHNSRGFGVAIVGNYYTAAALPTEAALRTVDRDTPSCAVRAGLLRPDYALLG 61  
Db 257 GWHIQGSHTYGFNDIALGIAFIGYFVEKPPNAAALEAAQD-LIQCAVVEGYLTPNYLLMG 315  
QY 62 HRQLVRTDCPGDALFDLLRTWPHF 85  
Db 316 HSDVNVILSPGQALYNIISTWPHF 339

RESULT 12  
ADE40166  
ID ADE40166 standard; protein; 341 AA.  
XX  
AC ADE40166;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human NOV26b protein - SEQ ID 72.  
XX  
KW NOVX; cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;  
KW antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;  
KW antiparkinsonian; antiasthmatic; gynaecological; cardiomyopathy;  
KW atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;

KW multiple sclerosis; graft-versus-host disease; Alzheimer's; Parkinson's;  
KW asthma; fertility disorder; vaccine; gene therapy; chromosome mapping;  
XX tissue typing; human; NOV.

OS Homo sapiens.

PN WO2003064589-A2.

PD 07-AUG-2003.

PF 02-AUG-2002; 2002WO-US024483.

PR 02-AUG-2001; 2001US-0309501P.

PR 03-AUG-2001; 2001US-0310291P.

PR 07-AUG-2001; 2001US-0310544P.

PR 08-AUG-2001; 2001US-0310951P.

PR 09-AUG-2001; 2001US-0311292P.

PR 13-AUG-2001; 2001US-0311979P.

PR 16-AUG-2001; 2001US-0312892P.

PR 17-AUG-2001; 2001US-0313201P.

PR 17-AUG-2001; 2001US-0313415P.

PR 20-AUG-2001; 2001US-0313643P.

PR 20-AUG-2001; 2001US-0313702P.

PR 21-AUG-2001; 2001US-0314031P.

PR 23-AUG-2001; 2001US-0314466P.

PR 28-AUG-2001; 2001US-0315403P.

PR 29-AUG-2001; 2001US-0315853P.

PR 17-SEP-2001; 2001US-0322716P.

PR 21-SEP-2001; 2001US-0323994P.

PR 14-DEC-2001; 2001US-0340233P.

PR 05-FEB-2002; 2002US-0354591P.

PR 19-MAR-2002; 2002US-0365478P.

PR 19-APR-2002; 2002US-0373814P.

PR 19-APR-2002; 2002US-0373825P.

PR 19-APR-2002; 2002US-0373989P.

PR 23-APR-2002; 2002US-0374632P.

PR 07-JUN-2002; 2002US-0386971P.

PR 01-AUG-2002; 2002US-00210172.

XX (CURA-) CURAGEN CORP.

PA Kekuda R, Miller CE, Patturajan M, Pena CEA, Rieger DK;

XX Shimkets RA, Zerhusen BD, Li L, Ji W, Padigaru M, Casman SJ;

PI Voss EZ, Boldog FL, Gorman L, Leite MW, Vernet CAM, Anderson DW;

PI Guo X, Zhong M, Gerlach VL, Hjal T, Rastelli L, Spytek KA;

PI Edinger SR, Ellerman K, Malyankar UM, Macdougall JR, Stone DJ;

PI Alsobrook JP, Lepley DM, Burgess CE, Majumder K, Wolenc AR;

PI Smithson G;

XX WPI; 2003-663472/62.

DR N-PSDB; ADE40165.

XX

PT New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.

XX Claim 1; SEQ ID NO 72; 560pp; English.

PS The invention relates to a novel NOVX polypeptide. The polypeptide of the  
XX invention demonstrates cardiant, antiarteriosclerotic, hypotensive,  
CC cytotstatic, anorectic, antidiabetic, immunosuppressive, anti-HIV,  
CC neuroprotective, nootropic, antiparkinsonian, antiasthmatic and  
CC gynaecological activities and may be useful in diagnosing, treating or  
CC preventing NOVX-associated disorders including cardiomyopathy,  
CC atherosclerosis, hypertension, cancer, obesity, diabetes, AIDS, multiple  
CC sclerosis, graft-versus-host disease, Alzheimer's disease, Parkinson's  
CC disease, asthma or fertility disorders. Furthermore, the polypeptides may  
CC be utilised as vaccines whilst the nucleic acids may be used as  
CC hybridisation probes, in gene therapy, chromosome mapping, tissue typing,  
CC preventative medicine and pharmacogenomics. The current sequence is that of  
CC the human NOV protein of the invention.

XX

SQ Sequence 341 AA;

Query Match

Best Local Similarity 21.1%; Score 198.5; DB 7; Length 341;

Matches 37; Conservative 12; Mismatches 34; Indels 1; Gaps 1;

QY 2 GWHVGAHTLGHNSRGFGVAIVGNYYTALPTEAALRTVRDTLPSCAVRAGLLRPDYALLG 61

Db 257 GWHIQGSHYVGFNDIALGIAFIGYFVEKFPNAAALEAAQD-LIQCAVVEGYLTFNYLLMG 315

QY 62 HRQLVRTDCPDALFDLLRTWPHF 85

Db 316 HSDVVNLSFGQALYNIISTWPHF 339

RESULT 13

ADE40164

ID ADE40164 standard; protein; 368 AA.

XX ADE40164;

AC ADE40164;

XX 29-JAN-2004 (first entry)

DT Human NOV26a protein - SEQ ID 70.

XX

DE NOVX; cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;  
KW antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic;  
KW antiparkinsonian; antiasthmatic; gynaecological; cardiomyopathy;  
KW atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS;  
KW multiple sclerosis; graft-versus-host disease; Alzheimer's; Parkinson's;  
KW asthma; fertility disorder; vaccine; gene therapy; chromosome mapping;  
KW tissue typing; human; NOV.

XX Homo sapiens.

OS WO2003064589-A2.

XX 07-AUG-2003.

XX 02-AUG-2002; 2002WO-US024483.

PF 02-AUG-2002; 2001US-0309501P.

XX 03-AUG-2001; 2001US-0310291P.

PR 07-AUG-2001; 2001US-0310544P.

PR 08-AUG-2001; 2001US-0310951P.

PR 09-AUG-2001; 2001US-0311292P.

PR 13-AUG-2001; 2001US-0311979P.

PR 16-AUG-2001; 2001US-0312892P.

PR 17-AUG-2001; 2001US-0313201P.

PR 17-AUG-2001; 2001US-0313415P.

PR 20-AUG-2001; 2001US-0313643P.

PR 20-AUG-2001; 2001US-0313702P.

PR 21-AUG-2001; 2001US-0314031P.

PR 23-AUG-2001; 2001US-0314466P.

PR 28-AUG-2001; 2001US-0315403P.

PR 29-AUG-2001; 2001US-0315853P.

PR 17-SEP-2001; 2001US-0322716P.

PR 21-SEP-2001; 2001US-0323994P.

PR 14-DEC-2001; 2001US-0340233P.

PR 05-FEB-2002; 2002US-0354591P.

PR 19-MAR-2002; 2002US-0365478P.

PR 19-APR-2002; 2002US-0373814P.

PR 19-APR-2002; 2002US-0373825P.

PR 19-APR-2002; 2002US-0373989P.

PR 23-APR-2002; 2002US-0374632P.

PR 07-JUN-2002; 2002US-0386971P.

PR 01-AUG-2002; 2002US-00210172.

XX (CURA-) CURAGEN CORP.

PA Kekuda R, Miller CE, Patturajan M, Pena CEA, Rieger DK;

XX Shimkets RA, Zerhusen BD, Li L, Ji W, Padigaru M, Casman SJ;

PI Voss EZ, Boldog FL, Gorman L, Leite MW, Vernet CAM, Anderson DW;







CC 2003 to standardise OS field)

XX

SQ Sequence 173 AA;

Query Match 20.8%; Score 195.5; DB 2; Length 173;  
Best Local Similarity 44.0%; Pred. No. 8e-12;  
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 2 GWHVGAHTLGHNSRGFGVAIVGNVTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALLG 61  
Db 81 GWHVGAHTYGYNSRSIGVAFIGNFNTDEPFGAMLEALRSLL-RCGVERGHLAGDYRVVA 139

QY 62 HRQLVRTDCPGDALFDLLRTWPHF 85  
Db 140 HRQLIASBSPGRKLYNQIRRWPEW 163

Search completed: May 18, 2004, 16:11:33  
Job time : 60 secs

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	195.5	20.8	173	3	US-08-928-917C-1	Sequence 1, Appli
2	195.5	20.8	173	3	US-08-928-917C-4	Sequence 4, Appli
3	195.5	20.8	173	4	US-09-374-678-1	Sequence 1, Appli
4	195.5	20.8	173	4	US-09-374-678-4	Sequence 4, Appli
5	195.5	20.8	196	3	US-08-928-917C-2	Sequence 2, Appli
6	195.5	20.8	196	3	US-08-928-917C-6	Sequence 6, Appli
7	195.5	20.8	196	3	US-08-928-917C-8	Sequence 8, Appli
8	195.5	20.8	196	4	US-09-374-678-2	Sequence 2, Appli
9	195.5	20.8	196	4	US-09-374-678-6	Sequence 6, Appli
10	195.5	20.8	196	4	US-09-374-678-8	Sequence 8, Appli
11	193	20.5	196	4	US-09-469-242-6	Sequence 6, Appli
12	183.5	19.5	368	4	US-09-469-242-4	Sequence 4, Appli
13	154	16.4	182	3	US-08-893-764-2	Sequence 2, Appli
14	151	16.0	190	3	US-08-799-149C-3	Sequence 3, Appli
15	106.5	11.3	957	4	US-09-252-991A-20408	Sequence 20408, A
16	99	10.5	243	4	US-09-469-242-2	Sequence 2, Appli
17	98	10.4	202	4	US-09-252-991A-32054	Sequence 32054, A
18	97.5	10.4	254	4	US-09-252-991A-31621	Sequence 31621, A
19	95.5	10.1	618	4	US-09-252-991A-29306	Sequence 29306, A
20	95	10.1	165	3	US-08-799-149C-5	Sequence 5, Appli
21	95	10.1	213	4	US-09-252-991A-22585	Sequence 22585, A
22	92.5	9.8	173	4	US-09-252-991A-19987	Sequence 19987, A
23	92.5	9.8	369	4	US-09-252-991A-29670	Sequence 29670, A
24	91	9.7	138	4	US-09-252-991A-20292	Sequence 20292, A
25	91	9.7	254	4	US-09-252-991A-20551	Sequence 20551, A
26	91	9.7	469	4	US-09-252-991A-25438	Sequence 25438, A
27	90.5	9.6	266	4	US-09-252-991A-32478	Sequence 32478, A





```
/
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 281/47701
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 523-3400
/ TELEFAX: (617) 523-6440
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 173 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-09-374-678-4

Query Match 20.8%; Score 195.5; DB 4; Length 173;
Best Local Similarity 44.0%; Pred. No. 4.4e-13;
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 2 GWHWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVDRDTLPSCAVRAGLLRPDYALLG 61
Db 81 GWLHVGAHTYGYNSRSIGVAFIGNFTDEPSGAMLEALRSLL-RCGVERGHLAGDYRVVA 139

QY 62 HRQLVRTDCPGDALEFDLLRTWPHF 85
Db 140 HRQLIASESPGRKLYNQIRRWPEW 163

RESULT 5
US-08-928-917C-2
; Sequence 2, Application US/08928917C
; Patent No. 6034217
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; APPLICANT: Ochiai, Masanori
; APPLICANT: Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; TITLE OF INVENTION: THEIR PRODUCTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109-4280
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,917C
; FILING DATE: 12-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 281/47701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-917C-2

Query Match 20.8%; Score 195.5; DB 3; Length 196;
Best Local Similarity 44.0%; Pred. No. 5.2e-13;
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 2 GWHWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVDRDTLPSCAVRAGLLRPDYALLG 61
Db 81 GWLHVGAHTYGYNSRSIGVAFIGNFTDEPSGAMLEALRSLL-RCGVERGHLAGDYRVVA 139

QY 62 HRQLVRTDCPGDALEFDLLRTWPHF 85
Db 140 HRQLIASESPGRKLYNQIRRWPEW 163

RESULT 6
US-08-928-917C-6
; Sequence 6, Application US/08928917C
; Patent No. 6034217
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; APPLICANT: Ochiai, Masanori
; APPLICANT: Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; TITLE OF INVENTION: THEIR PRODUCTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109-4280
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,917C
; FILING DATE: 12-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 281/47701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-917C-6

Query Match 20.8%; Score 195.5; DB 3; Length 196;
Best Local Similarity 44.0%; Pred. No. 5.2e-13;
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 2 GWHWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVDRDTLPSCAVRAGLLRPDYALLG 61
Db 104 GWLHVGAHTYGYNSRSIGVAFIGNFTDEPSGAMLEALRSLL-RCGVERGHLAGDYRVVA 162

QY 62 HRQLVRTDCPGDALEFDLLRTWPHF 85
Db 163 HRQLIASESPGRKLYNQIRRWPEW 186

RESULT 7
US-08-928-917C-8
; Sequence 8, Application US/08928917C
; Patent No. 6034217
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; APPLICANT: Ochiai, Masanori
; APPLICANT: Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; TITLE OF INVENTION: THEIR PRODUCTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
```

```
Db 104 GWLHVGAHTYGYNSRSIGVAFIGNFTDEPSGAMLEALRSLL-RCGVERGHLAGDYRVVA 162

QY 62 HRQLVRTDCPGDALEFDLLRTWPHF 85
Db 163 HRQLIASESPGRKLYNQIRRWPEW 186

RESULT 6
US-08-928-917C-6
; Sequence 6, Application US/08928917C
; Patent No. 6034217
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; APPLICANT: Ochiai, Masanori
; APPLICANT: Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; TITLE OF INVENTION: THEIR PRODUCTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109-4280
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,917C
; FILING DATE: 12-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 281/47701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-917C-6

Query Match 20.8%; Score 195.5; DB 3; Length 196;
Best Local Similarity 44.0%; Pred. No. 5.2e-13;
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 2 GWHWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVDRDTLPSCAVRAGLLRPDYALLG 61
Db 104 GWLHVGAHTYGYNSRSIGVAFIGNFTDEPSGAMLEALRSLL-RCGVERGHLAGDYRVVA 162

QY 62 HRQLVRTDCPGDALEFDLLRTWPHF 85
Db 163 HRQLIASESPGRKLYNQIRRWPEW 186

RESULT 7
US-08-928-917C-8
; Sequence 8, Application US/08928917C
; Patent No. 6034217
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; APPLICANT: Ochiai, Masanori
; APPLICANT: Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; TITLE OF INVENTION: THEIR PRODUCTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
```







ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/893,764  
FILING DATE: (Herewith)  
CLASSIFICATION: 515  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.1630000  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 182 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-893-764-2

Query Match 16.4%; Score 154; DB 3; Length 182;  
Best Local Similarity 36.8%; Pred. No. 1.2e-08;  
Matches 32; Conservative 16; Mismatches 35; Indels 4; Gaps 3;  
QY 1 RQWHWVGAHTLG--HNSRGFGVAIVGNYTAALPTEAALRTVTDLPSCAVRAGLLRPDYA 58  
Db 96 RGNWIKGDHT-GPIWNPMISGTFMGFMFMDRVPAKRALRAALNLL-ECGVSRGFLRSNYE 153  
QY 59 LLGHRQLVRTDCPGDALFDLLRTWPHF 85  
Db 154 VKGHRDVQSTLSPGDQLYQVQSWEHY 180

RESULT 14  
US-08-799-149C-3  
Sequence 3, Application US/08799149C  
Patent No. 6008195  
GENERAL INFORMATION:  
APPLICANT: Michael E. Selsted  
TITLE OF INVENTION: Antimicrobial Peptides and  
TITLE OF INVENTION: Methods of Use  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/799,149C  
FILING DATE: 14-February-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/011,834  
FILING DATE: 16-February-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lisa A. Haile, Ph.D.  
REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07306/009001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 39..598  
US-08-799-149C-3

Query Match 16.0%; Score 151; DB 3; Length 190;  
Best Local Similarity 34.8%; Pred. No. 2.6e-08;  
Matches 31; Conservative 18; Mismatches 38; Indels 2; Gaps 2;  
QY 1 RQWHWVGAHT-LGHNSRGFGVAIVGNYTAALPTEAALRTVTDLPSCAVRAGLLRPDYAL 59  
Db 103 RGNWTKGDHSGPTWNPFAIGISFGMGNVHRVFFASALRAAQSL-ACGAARGVLTENYEV 161  
QY 60 LQHRQLVRTDCPGDALFDLLRTWPHFTAV 88  
Db 162 KGHDRVQQTLSPGDELYKIIQQWPHYRRV 190

RESULT 15  
US-09-252-991A-20408  
Sequence 20408, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 20408  
LENGTH: 957  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20408

Query Match 11.3%; Score 106.5; DB 4; Length 957;  
Best Local Similarity 29.4%; Pred. No. 0.012;  
Matches 64; Conservative 14; Mismatches 69; Indels 71; Gaps 12;  
QY 8 AHTLGHNSRGFG-----VAIVG--NYTAALPTEAALRTVTDLPSCAVRAGLLRPDY 57  
Db 3 AHTLGHNSIRRSADGQPLEPPPAAVGPGAPASLES-TARRTGRS--PAADRRARRIRQEF 59  
QY 58 ALLGHRQLVRTDCPGDALFDLL-----RTWPH----- 84  
Db 60 P--GHRVLRK---PRSAALAEPLAGTEQPRRSRTFFRTPARWLATVPDPGRRGRPAED 114  
QY 85 ---FTAVSLRSLHYTARRPSPVYTSSTRPLPPACNSCARTASARPTSRHV-----Y 133  
Db 115 APASPAVRLRDL---ARRPA---RRTRPLPGPAASPATGAGRLPPGPGRGARPLPAIPPQ 168  
QY 134 SGNLGPFAFAGH---SAGNIPDPVTSAYAASAQPTQPA 168  
Db 169 SPARGPGAAGHQPAAGLAPGAPAPVAIAIARTQTGPA 206

Search completed: May 18, 2004, 16:14:01  
Job time : 23 secs



---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:12:39 ; Search time 43 Seconds  
(without alignments)  
1125.988 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 941  
Sequence: 1 RGHWVGAHTLGHNSRGFGV.....SAYAASAPQTQACFPSS 174

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	941	100.0	174	15	US-10-068-956-2
2	941	100.0	634	14	US-10-239-663-63
3	941	100.0	634	15	US-10-094-749-2263
4	483.5	51.4	576	14	US-10-239-663-64
5	483.5	51.4	576	16	US-10-467-248-9
6	403	42.8	530	15	US-10-068-956-4
7	198.5	21.1	341	12	US-10-210-172-72
8	198.5	21.1	368	12	US-10-210-172-70
9	195.5	20.8	173	14	US-10-135-207-1
10	195.5	20.8	173	14	US-10-135-207-4
11	195.5	20.8	196	14	US-10-135-207-2
12	195.5	20.8	196	14	US-10-135-207-6
13	195.5	20.8	196	14	US-10-135-207-8
14	193	20.5	196	10	US-09-984-130-36
15	193	20.5	196	10	US-09-946-374-216

16	193	20.5	196	10	US-09-836-353A-36	Sequence 36, Appl
17	193	20.5	196	12	US-10-006-485A-216	Sequence 216, App
18	193	20.5	196	12	US-10-013-907A-216	Sequence 216, App
19	193	20.5	196	12	US-10-015-499A-216	Sequence 216, App
20	193	20.5	196	12	US-10-013-910A-216	Sequence 216, App
21	193	20.5	196	12	US-10-226-254A-216	Sequence 216, App
22	193	20.5	196	12	US-10-015-395A-216	Sequence 216, App
23	193	20.5	196	14	US-10-006-856A-216	Sequence 216, App
24	193	20.5	196	14	US-10-006-818A-216	Sequence 216, App
25	193	20.5	196	14	US-10-015-393A-216	Sequence 216, App
26	193	20.5	196	14	US-10-015-869A-216	Sequence 216, App
27	193	20.5	196	14	US-10-012-121A-216	Sequence 216, App
28	193	20.5	196	14	US-10-006-116A-216	Sequence 216, App
29	193	20.5	196	14	US-10-006-117A-216	Sequence 216, App
30	193	20.5	196	14	US-10-017-527A-216	Sequence 216, App
31	193	20.5	196	14	US-10-013-913A-216	Sequence 216, App
32	193	20.5	196	14	US-10-007-194A-216	Sequence 216, App
33	193	20.5	196	14	US-10-013-430A-216	Sequence 216, App
34	193	20.5	196	14	US-10-011-671A-216	Sequence 216, App
35	193	20.5	196	14	US-10-012-755A-216	Sequence 216, App
36	193	20.5	196	14	US-10-015-386A-216	Sequence 216, App
37	193	20.5	196	14	US-10-011-692A-216	Sequence 216, App
38	193	20.5	196	14	US-10-006-768A-216	Sequence 216, App
39	193	20.5	196	14	US-10-017-610A-216	Sequence 216, App
40	193	20.5	196	14	US-10-006-063A-216	Sequence 216, App
41	193	20.5	196	14	US-10-020-063A-216	Sequence 216, App
42	193	20.5	196	14	US-10-015-391A-216	Sequence 216, App
43	193	20.5	196	14	US-10-017-407A-216	Sequence 216, App
44	193	20.5	196	14	US-10-011-833A-216	Sequence 216, App
45	193	20.5	196	14	US-10-006-041A-216	Sequence 216, App

ALIGNMENTS

RESULT 1  
US-10-068-956-2  
Sequence 2, Application US/10068956  
Publication No. US20030204065A1

GENERAL INFORMATION:

; APPLICANT: Paul Young et al.  
; TITLE OF INVENTION: PGRP-L Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PF513P1  
; CURRENT APPLICATION NUMBER: US/10/068,956  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: 60/149,715  
; PRIOR FILING DATE: 1999-08-20  
; PRIOR APPLICATION NUMBER: PCT/US00/22877  
; PRIOR FILING DATE: 2000-08-18  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: human  
US-10-068-956-2

*Original*

Query Match	100.0%;	Score 941;	DB 15;	Length 174;
Best Local Similarity	100.0%;	Pred. No. 1.8e-83;		
Matches 174;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	RGHWVGAHTLGHNSRGFGV	1	RGHWVGAHTLGHNSRGFGV
Db	1	RGHWVGAHTLGHNSRGFGV	1	RGHWVGAHTLGHNSRGFGV
QY	61	GHRLVTRDCGDALFDLLRTPHFTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART	120	
Db	61	GHRLVTRDCGDALFDLLRTPHFTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART	120	
QY	121	ASARPPTSRRHVSGNLGPAFAGHSAGNIPDPVTSAYAASAPQTQACFPSS	174	
Db	121	ASARPPTSRRHVSGNLGPAFAGHSAGNIPDPVTSAYAASAPQTQACFPSS	174	



```
RESULT 2
US-10-239-663-63
; Sequence 63, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-663-63

Query Match      100.0%; Score 941; DB 14; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.5e-83;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60
Db      461 RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520

QY      61  GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 120
Db      521 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 580

QY      121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAQAQPTQACPFSS 174
Db      581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAQAQPTQACPFSS 634

RESULT 3
US-10-094-749-2263
; Sequence 2263, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA

Query Match      100.0%; Score 941; DB 14; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.5e-83;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60
Db      461 RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520

QY      61  GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 120
Db      521 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 580

QY      121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAQAQPTQACPFSS 174
Db      581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAQAQPTQACPFSS 634
```

```
FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2263
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2263

Query Match      100.0%; Score 941; DB 15; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.5e-83;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60
Db      461 RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520

QY      61  GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 120
Db      521 GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 580

QY      121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAQAQPTQACPFSS 174
Db      581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAQAQPTQACPFSS 634

RESULT 4
US-10-239-663-64
; Sequence 64, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-663-64

Query Match      51.4%; Score 483.5; DB 14; Length 576;
Best Local Similarity 83.3%; Pred. No. 2.2e-38;
Matches 95; Conservative 0; Mismatches 10; Indels 9; Gaps 2;

QY      1  RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60
Db      461 RGWHWVGAAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520

QY      61  GHRQLVRTDCPGDALFDLLRTWPHTAVSLRSLHYTARRP--SVYTSSTRPLPP 112
```

Db 521 GHRQLVRTDCPGDALFDLLRTPHFTAT-----VKPRPARSVSKRSRREPPPP 567

RESULT 5

US-10-467-248-9 Application US/10467248  
; Sequence 9, Application US/10467248  
; Publication No. US20040086905A1  
; GENERAL INFORMATION:  
; APPLICANT: DAS, Debopriya; YAO, Monique G.;  
; APPLICANT: ARVIZU, Chandra S.; BAUGHN, Mariah R.;  
; APPLICANT: LU, Yan; HAFALIM, April J.A.;  
; APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;  
; APPLICANT: LU, Dyung Aina M.; YUE, Henry;  
; APPLICANT: DING, Li; ELLIOTT, Vicki S.;  
; APPLICANT: FORSYTHE, Ian J.; RAMKUMAR, Jayalaxmi;  
; APPLICANT: GANDHI, Ameena R.; ISON, Craig H.;  
; APPLICANT: WARREN, Bridget A.; TANG, Y. Tom;  
; APPLICANT: EMERLING, Brooke M.; HONCHELL, Cynthia D.;  
; APPLICANT: LYNE, Michael; BARROSO, Ines  
; TITLE OF INVENTION: LIPID-ASSOCIATED MOLECULES  
; FILE REFERENCE: PI-0358 USN  
; CURRENT APPLICATION NUMBER: US/10/467,248  
; CURRENT FILING DATE: 2003-08-06  
; PRIOR APPLICATION NUMBER: PCT/US02/03813  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: US 60/266,910  
; PRIOR FILING DATE: 2001-02-06  
; PRIOR APPLICATION NUMBER: US 60/276,891  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: US 60/276,855  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: US 60/279,760  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 60/283,818  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/285,405  
; PRIOR FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 576  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 6897474CD1

US-10-467-248-9

Query Match 51.4%; Score 483.5; DB 16; Length 576;  
Best Local Similarity 83.3%; Pred. No. 2.2e-38;  
Matches 95; Conservative 0; Mismatches 10; Indels 9; Gaps 2;

Qy 1 RGWHVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVVRTDLPSCAVRAGLLRPDYALL 60  
Db 461 RGWHVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVVRTDLPSCAVRAGLLRPDYALL 520  
Qy 61 GHRQLVRTDCPGDALFDLLRTPHFTAVSLRSLHYTARRP--SVYTSSTRPLPP 112  
Db 521 GHRQLVRTDCPGDALFDLLRTPHFTAT-----VKPRPARSVSKRSRREPPPP 567

RESULT 6

US-10-068-956-4  
; Sequence 4, Application US/10068956  
; Publication No. US20030204065A1  
; GENERAL INFORMATION:  
; APPLICANT: Paul Young et al.  
; TITLE OF INVENTION: PGRP-L Polynucleotides, polypeptides, and Antibodies  
; FILE REFERENCE: PF513P1  
; CURRENT APPLICATION NUMBER: US/10/068,956  
; CURRENT FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: 60/149,715  
; PRIOR FILING DATE: 1999-08-20

022-052X

; PRIOR APPLICATION NUMBER: PCT/US00/22877  
; PRIOR FILING DATE: 2000-08-18  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 530  
; TYPE: PRT  
; ORGANISM: human  
US-10-068-956-4

Query Match 42.8%; Score 403; DB 15; Length 530;  
Best Local Similarity 81.8%; Pred. No. 1.4e-30;  
Matches 72; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 1 RGWHVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVVRTDLPSCAVRAGLLRPDYALL 60  
Db 441 RGWHVGAHTRGYNSRGFGVAFVGNYYTGSLENEAALNTVRDALPSCAIREGLLRPDYKLL 500  
Qy 61 GHRQLVRTDCPGDALFDLLRTPHFTAV 88  
Db 501 GHRQLVLTCHCPGNALFNLLRTPHFTEV 528

RESULT 7

US-10-210-172-72  
; Sequence 72, Application US/10210172  
; Publication No. US20040043928A1  
; GENERAL INFORMATION:

; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Miller, Charles  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Pena, Carol  
; APPLICANT: Rieger, Daniel  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Li, Li  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Casman, Stacie  
; APPLICANT: Voss, Edward  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Gorman, Linda  
; APPLICANT: Leite, Mario  
; APPLICANT: Vernet, Corine  
; APPLICANT: Anderson, David  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Zhong, Mei  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Hjalt, Tord  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David  
; APPLICANT: Alsobrook II, John  
; APPLICANT: Lepley, Denise et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-416 A  
; CURRENT APPLICATION NUMBER: US/10/210,172  
; CURRENT FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/309,501  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: 60/323,994  
; PRIOR FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/373,814  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/310,291  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: 60/310,951  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: 60/310,544





[illegible]

```

RESULT 10
US-10-135-207-4
; Sequence 4, Application US/10135207
; Publication No. US20030166068A1
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; Ochiai, Masanori
; Tsuchiya, Masakazu
; TITLE OF INVENTION: PETIDOLYCAN RECOGNITION PROTEINS AND
; THEIR PRODUCTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109-4280
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/135,207
; FILING DATE: 30-Apr-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,917C
; FILING DATE: 12-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 281/47701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 173 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-135-207-4

```

```

QY      2  GWHWVGAKTLGHSRGVGVVGNVYTAALPTEAALRTVVRTDILPSCAVRAGLLRPDYALLG 61
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      81  GWLHVGAHTYGYNSRSIGVAFIGNFNTDEPSGAMLEALRSL-RCGVERGHLAGDYRVVA 139
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      62  HRQLVRTDCPGDALFDLLRTWPHF 85
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      140 HRQLIASESPGRKLYNQIRRWPEW 163
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
US-10-135-207-2
; Sequence 2, Application US/10135207
; Publication No. US20030166068A1
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; Ochiai, Masanori
; Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; THEIR PRODUCTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109-4280
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/135,207
; FILING DATE: 30-Apr-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,917C
; FILING DATE: 12-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 281/47701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-135-207-2

```

```
Query Match      20.8%; Score 195.5; DB 14; Length 196;  
Best Local Similarity 44.0%; Pred. No. 6.1e-11;  
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1  
  
QY    2   GWHWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRBDTLPSCAVRAGLLRPDYALLG 61  
       ||| |||| |::||| :||| ::||| :||| :||| :||| :|||  
Db     104 GWLHVGAHTYGYNRSRISGVAFIGNFTDEPSGMLEALSLL-RCGVERGHLAGD YRVVA 162  
                                               |:::  
  
QY    62 HRQLVRTDCPGDALFDLLRTWPHF 85  
       |||:::||| ::||| :||| :  
Db     163 HRQLIASESPGRKLYNQIRRWPEW 186  
                               |||::|||  
  
RESULT 12  
US-10-135-207-6  
; Sequence 6, Application US/10135207  
; Publication No. US20030166068A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashida, Masaaki
```



; Ochiai, Masanori
; Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; THEIR PRODUCTION
;
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109-4280
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/135,207
; FILING DATE: 30-Apr-2002
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,917C
; FILING DATE: 12-SEP-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 281/47701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-135-207-6

Query Match 20.8%; Score 195.5; DB 14; Length 196;
Best Local Similarity 44.0%; Pred. No. 6.1e-11;
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

Qy 2 GWHVVGHAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVVRTDLPSCAVRAGLLRPDYALLG 61
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
104 GWLHVGAHTYGYNSRSIGVAFIGNFTDEPSGAMLEALRSLL-RCGVERGHLAGDYRVVA 162
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 62 HRQLVRTDCPGDALFDLLRTWPHF 85
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 163 HRQLIASESPGRKLYNQIRRWPEW 186
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 13
US-10-135-207-8
; Sequence 8, Application US/10135207
; Publication No. US20030166068A1
; GENERAL INFORMATION:
; APPLICANT: Ashida, Masaaki
; Ochiai, Masanori
; Tsuchiya, Masakazu
; TITLE OF INVENTION: PEPTIDOGLYCAN RECOGNITION PROTEINS AND
; THEIR PRODUCTION
;
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109-4280
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/135,207
; FILING DATE: 30-Apr-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,917C
; FILING DATE: 12-SEP-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 281/47701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 196 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-135-207-8

Query Match 20.8%; Score 195.5; DB 14; Length 196;
Best Local Similarity 44.0%; Pred. No. 6.1e-11;
Matches 37; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

Qy 2 GWHVVGHAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVVRTDLPSCAVRAGLLRPDYALLG 61
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
104 GWLHVGAHTYGYNSRSIGVAFIGNFTDEPSGAMLEALRSLL-RCGVERGHLAGDYRVVA 162
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 62 HRQLVRTDCPGDALFDLLRTWPHF 85
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 163 HRQLIASESPGRKLYNQIRRWPEW 186
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 14
US-09-984-130-36
; Sequence 36, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-36

Query Match 20.5%; Score 193; DB 10; Length 196;
Best Local Similarity 42.5%; Pred. No. 1.1e-10;
Matches 37; Conservative 17; Mismatches 29; Indels 4; Gaps 3;

Qy 1 RGWHVVGHAHTLGH--NSRGFGVAIVGNYTAALPTEAALRTVVRTDLPSCAVRAGLLRPDYA 58
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 109 RGNWFTGAHS-GHLWNFMSIGISFMGNVMDRVPTFPQAIRAAQGLL-ACGVAQGAIRSNYV 166
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 59 LLGHRQLVRTDCPGDALFDLLRTWPHF 85  
Db 167 LKGRDVGQRTLSPGNQLYHLIQNWPHY 193

RESULT 15

US-09-946-374-216  
; Sequence 216, Application US/09946374  
; Publication No. US20030073129A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2830P1C1  
; CURRENT APPLICATION NUMBER: US/09/946,374  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/098716  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: 60/098723  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: 60/098749  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: 60/098750  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: 60/098803  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR APPLICATION NUMBER: 60/098821  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR APPLICATION NUMBER: 60/098843  
; PRIOR FILING DATE: 1998-09-02  
; PRIOR APPLICATION NUMBER: 60/099536  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099596  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099598  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099602  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099642  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 60/099741  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099754  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099763  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099792  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099808  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099812  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099815

; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/099816  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/100385  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100388  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100390  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/100584  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100627  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100661  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100662  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100664  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: 60/100683  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/100684  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/100710  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/100711  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/100848  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/100849  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/100919  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/100930  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: 60/101014  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/101068  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/101071  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/101279  
; PRIOR FILING DATE: 1998-09-22  
; PRIOR APPLICATION NUMBER: 60/101471  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101472  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101474  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101475  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101476  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101477  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101479  
; PRIOR FILING DATE: 1998-09-23  
; PRIOR APPLICATION NUMBER: 60/101738  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101741  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101743  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101915  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/101916  
; PRIOR FILING DATE: 1998-09-24  
; PRIOR APPLICATION NUMBER: 60/102207  
; PRIOR FILING DATE: 1998-09-29  
; PRIOR APPLICATION NUMBER: 60/102240  
; PRIOR FILING DATE: 1998-09-29  
; PRIOR APPLICATION NUMBER: 60/102307  
; PRIOR FILING DATE: 1998-09-29











QY 121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQIP 167  
Db 174 PGVRRVTPPRHRFSGQ-GAAYGGRQAQALP-LLQQGASENLKPPSRP 218

RESULT 7  
T49726  
hypothetical protein B23L21.390 [imported] - Neurospora crassa  
C;Species: Neurospora crassa  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C;Accession: T49726  
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A;Reference number: Z25022  
A;Accession: T49726  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1198 <SCH>  
A;Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.390  
A;Experimental source: BAC clone B23L21; strain OR74A  
C;Genetics:  
A;Gene: NCSP:B23L21.390  
A;Map position: 6  
A;Introns: 10/3; 188/3

Query Match 9.2%; Score 87; DB 2; Length 1198;  
Best Local Similarity 26.2%; Pred. No. 9.2;  
Matches 43; Conservative 19; Mismatches 74; Indels 28; Gaps 5;

QY 28 AALPTEAALRTVRDTLPSCAVRAGLLRPD---YALLGHRQLVRTDCP----- 71  
Db 780 AELPAELPVEISSEQLP--PARAGSVTSDYLEYMDGSPSSRRTDAPPKTRTGSVSSQGS 837

QY 72 -GDALFDLLRTWPHFTAVSLRSLHYTARRPVSVYTSSTRPLPPACNSCARTASARPPPTSRR 130  
Db 838 SGAATQETKSSGSHSRQGSAGSVSKAAGAGNSTSSRK-----GSAVSSVSSSSS 889

QY 131 HVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQACPFPPSS 174  
Db 890 HQRQGSVGSVASTDSSRRDPE-TTTHARQASNDSGAATTPAS 932

RESULT 8  
T36390  
hypothetical protein SCE94.30c - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T36390  
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, April 1999  
A;Reference number: Z21573  
A;Accession: T36390  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-208 <OLI>  
A;Cross-references: EMBL:AL049628; PIDN:CAB40879.1; GSPDB:GN00070; SCOEDB:SCE94.30c  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCOEDB:SCE94.30c

Query Match 9.1%; Score 85.5; DB 2; Length 208;  
Best Local Similarity 30.6%; Pred. No. 1.8;  
Matches 49; Conservative 7; Mismatches 63; Indels 41; Gaps 8;

QY 12 GHNSRGFGVAIVGNVTAALPTEAALRTVRDTLPSCAVRAGLLRPDYAL-LGHRQLVRTDC 70  
Db 56 GARLRAFLDLADLARYRALPGTAT-----GYLALLDEALGAGH----RPDA 97

QY 71 PGDALFDLLRTWPHFTAVSLRSLHYTARRPVSVYTSSTRPLPPACNSCARTASARPPTSRR 130  
Db 98 DDLAALGALRGTPAAALLARCAPET---PTAPTASTVLPAA---RART-----RL 143

QY 131 HVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQIPACP 170

Db 144 HAVPG-----GROSGKEP-PQKPAAKPSRPPAPAPATP 175

RESULT 9  
AF3615  
glycosyl transferase (EC 2.4.1.1-) [imported] - Brucella melitensis (strain 16M)  
C;Species: Brucella melitensis  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C;Accession: AF3615  
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A;Reference number: AD3252; PMID:11756688  
A;Accession: AF3615  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-438 <KUR>  
A;Cross-references: GB:AE008918; PIDN:AAL54089.1; PID:gl7985047; GSPDB:GN00191  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEII0847  
A;Map position: 11  
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 9.1%; Score 85.5; DB 2; Length 438;  
Best Local Similarity 31.1%; Pred. No. 4.2;  
Matches 41; Conservative 11; Mismatches 49; Indels 31; Gaps 8;

QY 27 TAALPTE----AALRTVRDTLPSCAVRAGLLRPDYALLGH-----RQLVRTDCPG 72  
Db 137 SAALPIRFRRLVNVWMDLFPETAMELGLIRPD-TVSGRLAMALRDWSMRQSALTICPI 195

QY 73 DALFDLL--RTWPHFTAVSLRSLHYTARRPVSVYTSSTRPLPPACNSCARTASARPPTSRR 130  
Db 196 DRMAHYLAKRDIP---AESLGVVHHWADRNEIV-----PVEPVENPLRAWG----LGRK 243

QY 131 HV--YSGNLGPA 140  
Db 244 FVIGYSGNFGRA 255

RESULT 10  
S55316  
mucin (clone PGM-2B) - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 17-Nov-2000  
C;Accession: S55316  
R;Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; LaMont, J.T.  
Biochem. J. 308, 89-96, 1995  
A;Title: Isolation and characterization of cDNA clones encoding pig gastric mucin.  
A;Reference number: S55315; MUID:95275264; PMID:7755593  
A;Accession: S55316  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-317 <TUR>  
A;Cross-references: GB:U12768; NID:g915204; PIDN:AAC48525.1; PID:g915207  
C;Superfamily: pig submaxillary mucin

Query Match 9.0%; Score 85; DB 2; Length 317;  
Best Local Similarity 29.7%; Pred. No. 3.2;  
Matches 35; Conservative 13; Mismatches 40; Indels 30; Gaps 5;

QY 67 RTDCPGDALFDLLRTWPHFTAVSLRSLHYTARRPVSVYTSSTRPLPPACNSCARTASA--- 123  
Db 30 KKDCPVSPI-----TLPTTTSVRVTSPPPTSSHGA--TSSTTSVQPPSSSSSAPTTSATSV 82

QY 124 -----RPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQACPFPPSS 174  
Db 83 QPSSSGSAPTTSATSATSVQSSSSGSA-----PTTS--ATSVQPPSSSSSPPISST 127





Db	287	WLGLALVRLVFMALGTPAVGLLSDVLGALPAGLPORFATDPPLRSTMLAATLCATATLML	346
QY	48	-----VRAGLLRPDYALLGHRQLVRTDCPGDALFDLLRTWPHFTAVSLRSLHYTAR	98
Db	347	LTLVPLASSVNAGLRR-----RLWPNAAHPLGLAQAH----	377
QY	99	RPSVYTSSSTRPLPPACNSCARTASARPPTSRHRHVYSGNLGPAFAGHSAGNIPDPVTSAYA	158
Db	378	-----RQAAARQYAPRPAAAAAAGP----HQAG-----TYA	405
QY	159	ASAQPQTQPACPFPS	173
Db	406	ASATPAPAPARPAPS	420

RESULT 15  
T46292  
hypothetical protein DKFZp434E0610.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 18-Aug-2000  
C;Accession: T46292  
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23035  
A;Accession: T46292  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-616 <AAA>  
A;Cross-references: EMBL:AL137582  
A;Experimental source: adult testis; clone DKFZp434E0610  
C;Genetics:  
A;Note: DKFZp434E0610.1  
C;Superfamily: WW repeat homology  
F;62-100/Domain: WW repeat homology <WWR>

Query Match	8.8%	Score 82.5;	DB 2;	Length 616;	
Best Local Similarity	23.0%	Pred. No. 11;			
Matches	46;	Conservative 14;	Mismatches 83;	Indels 57;	Gaps 7;
QY	12	GHNSRGFGVAIVGNVYTAALPTEAALRTVRDTLPSCAVR--AGLLRPDYALLGHRQLVRTD	69		
Db	7	GHVGGGEG-----PTAAARPETRRPEPAPRTRAPAGRQPQPSMSAATHSPMMQVA	55		
QY	70	C-----PG-DALEDLLRTWPHFTAVSLRSLHYTARRPSPVYTSSTRPLPPACNSCAR	119		
Db	56	SGNGDRDPLPPGWEIKIDPQTGWPF--VDHNSRTTWNDRVPSEGPKEPSSANGPSR	113		
QY	120	TASARPPTSRHRHVYSGNLGPAFAGHSAGNIPDPV-----T	154		
Db	114	EGSRLPPAREGH-----PVYPQLRPGYIPIPVLHEGAENRQVHPFHVYPQGMQRFRT	166		
QY	155	SAYAASAQPQTQPACPFPS	174		
Db	167	EAATAAPQRSQSPLRGMPET	186		

Search completed: May 18, 2004, 16:12:35  
Job time : 23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:05:53 ; Search time 17 Seconds  
(without alignments)  
532.953 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 941  
Sequence: 1 RGWHVVGHTLGHNSRGFV.....SAYAASAPQTQPACFPFPSS 174

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483.5	51.4	576	1	PGPL_HUMAN
2	416.5	44.3	598	1	PGPL_PIG
3	408	43.4	530	1	PGPL_MOUSE
4	198.5	21.1	341	1	PGIA_HUMAN
5	196.5	20.9	182	1	PGRP_TRINI
6	193.5	20.6	196	1	PGRP_BOMMO
7	193	20.5	196	1	PGRP_HUMAN
8	183.5	19.5	373	1	PGIB_HUMAN
9	172	18.3	190	1	PGRP_BOVIN
10	167	17.7	193	1	PGRP_CAMDR
11	155	16.5	183	1	PGRP_RAT
12	154	16.4	182	1	PGRP_MOUSE
13	101	10.7	215	1	Y215_ADE02
14	84	8.9	682	1	HTF4_HUMAN
15	83.5	8.9	465	1	FXD1_HUMAN
16	82	8.7	659	1	TRM1_HUMAN
17	81.5	8.7	116	1	YE94_HAEIN
18	81.5	8.7	519	1	ERR1_HUMAN
19	81.5	8.7	712	1	IRAI_HUMAN
20	81	8.6	1109	1	TCF8_RAT
21	81	8.6	1914	1	STCK_EMBENI
22	80.5	8.6	2716	1	OSA_DROME
23	80	8.5	1117	1	TCF8_MOUSE
24	79.5	8.4	1887	1	RPB1_DROME
25	79	8.4	150	1	NAAA_BPT7
26	79	8.4	775	1	POPI_SCHPO
27	78.5	8.3	625	1	DUS8_HUMAN
28	78.5	8.3	710	1	IRAI_MOUSE
29	78.5	8.3	977	1	Y848_HUMAN
30	78.5	8.3	1043	1	TCF8_MESAU
31	78	8.3	1208	1	PER_DROYA
32	77.5	8.2	354	1	CD68_HUMAN
33	77.5	8.2	642	1	Z398_HUMAN

34	77	8.2	267	1	EXTN_MAIZE	P14918	zea mays (m
35	77	8.2	806	1	MK07_MOUSE	Q9wvs8	mus musculus
36	77	8.2	870	1	DYN2_HUMAN	P50570	homo sapien
37	77	8.2	2505	1	CCAA_HUMAN	O00555	homo sapien
38	76.5	8.1	732	1	POK_DROME	Q01842	drosophila
39	76.5	8.1	3421	1	TEGU_HSVB	P28955	equine herp
40	76	8.1	508	1	EGR1_RAT	P08154	rattus norv
41	76	8.1	802	1	ENAH_MOUSE	Q03173	mus musculus
42	75.5	8.0	530	1	NLFA_MOUSE	Q8bg30	mus musculus
43	75.5	8.0	928	1	RSC1_YEAST	P53236	saccharomyc
44	75.5	8.0	5703	1	MU5B_HUMAN	Q9hc84	homo sapien
45	75	8.0	150	1	NAAA_BPT3	P20331	bacterioph

ALIGNMENTS

RESULT 1  
PGPL\_HUMAN  
ID PGPL\_HUMAN STANDARD; PRT; 576 AA.  
AC Q96PD5; Q96N74;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE N-acetylmuramoyl-L-alanine amidase precursor (EC 3.5.1.28)  
DE (Peptidoglycan recognition protein long) (PGRP-L) (UNQ3103/PRO10102).  
GN PGRPL OR PGLYRPL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE DISTRIBUTION.  
RX MEDLINE=21433985; PubMed=11461926;  
RA Liu C., Xu Z., Gupta D., Dziarski R.;  
RT "Peptidoglycan recognition proteins: a novel family of four human  
RT innate immunity pattern recognition molecules.";  
RL J. Biol. Chem. 276:34686-34694(2001).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2),  
RC TISSUE=Liver;  
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,  
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,  
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,  
RA Masuho Y., Nagai K., Isogai T.;  
RT "NEDO human cDNA sequencing project.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=22887296; PubMed=12975309;  
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.-H., Yansura D.,  
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
RA Godowski P.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale  
RT effort to identify novel human secreted and transmembrane proteins: a  
RT bioinformatics assessment.";  
RL Genome Res. 13:2265-2270(2003).  
RN [4]  
RP SEQUENCE OF 22-36.  
RX MEDLINE=95392215; PubMed=74631175;  
RA De Pauw P., Neyt C., Vanderwinkel E., Wattiez R., Falmagne P.;  
RT "Characterization of human serum N-acetylmuramyl-L-alanine amidase  
RT purified by affinity chromatography.";

RL Protein Expr. Purif. 6:371-378(1995).

RN [5]

RP FUNCTION, AND MUTAGENESIS.

RX PubMed=14506276;

RA Wang Z.-M., Li X., Cocklin R.R., Wang M., Wang M., Fukase K.,

RA Inamura S., Kusumoto S., Gupta D., Dziarski R.;

RT "Human peptidoglycan recognition protein-L (PGRP-L) is an N-

RT acetylmuramoyl-L-alanine amidase.";

RL J. Biol. Chem. 278:10-0(2003).

RN [6]

RP CARBOHYDRATE-LINKAGE SITE ASN-485.

RX MEDLINE=22660472; PubMed=12754519;

RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;

RT "Identification and quantification of N-linked glycoproteins using

RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";

RL Nat. Biotechnol. 21:660-666(2003).

CC -!- FUNCTION: May plays a scavenger role by digesting biologically

CC active peptidoglycan (PGN) into biologically inactive fragments.

CC Has no direct bacteriolytic activity.

CC -!- CATALYTIC ACTIVITY: Hydrolyzes the link between N-acetylmuramoyl

CC residues and L-amino acid residues in certain bacterial cell-wall

CC glycopeptides.

CC -!- COFACTOR: Zinc (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted and membrane-associated.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=Q96PD5-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q96PD5-2; Sequence=VSP\_008964;

CC Note=May be due to an intron retention. No experimental

CC confirmation available;

CC -!- TISSUE SPECIFICITY: Strongly expressed in liver and fetal liver.

CC Expressed to a much lesser extent in traverse colon, lymph nodes,

CC heart, thymus, pancreas, descending colon, stomach and testis.

CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase

CC family 2.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF384856; AAL05629.1; -.

DR EMBL; AK055882; BAB71034.1; -.

DR EMBL; AY358156; AAQ88523.1; -.

DR MIM; 608199; -.

DR InterPro; IPR002502; Amidase\_2.

DR InterPro; IPR006619; PGRP.

DR Pfam; PF01510; Amidase\_2; 1.

DR SMART; SM00644; Ami\_2; 1.

DR SMART; SM00701; PGRP; 1.

KW Hydrolase; Immune response; Metal-binding; Zinc; Signal; Glycoprotein;

KW Alternative splicing.

FT SIGNAL 1 21

FT CHAIN 22 576 N-ACETYLMURAMOYL-L-ALANINE AMIDASE.

FT METAL 411 411 ZINC (BY SIMILARITY).

FT METAL 447 447 ZINC (BY SIMILARITY).

FT METAL 522 522 ZINC (BY SIMILARITY).

FT METAL 530 530 ZINC (BY SIMILARITY).

FT DISULFID 419 425 BY SIMILARITY.

FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .).

FT VARSPIC 548 576 TVKPRPARSVSKRSRREPPRTLPTADLQ -> VSLRSLHY

FT TARRPSVYTSRTPLPACNSCARTASARPPTSRHHVSGN

FT LGPAFAGHSAGNIPDPVTSAYASAQPTQACPPSS

FT (in isoform 2).

FT /FTId=VSP\_008964.

FT H->A: NO EFFECT ON AMIDASE ACTIVITY.

FT MUTAGEN 411 411

FT MUTAGEN 419 419 C->A: ABOLISHES AMIDASE ACTIVITY.

FT MUTAGEN 436 436 H->A: NO EFFECT ON AMIDASE ACTIVITY.

FT MUTAGEN 442 442 W->A: REDUCED AMIDASE ACTIVITY.

FT MUTAGEN 447 447 Y->A: ABOLISHES AMIDASE ACTIVITY.

FT MUTAGEN 530 530 C->S: ABOLISHES AMIDASE ACTIVITY.

FT CONFLICT 46 46 T -> A (IN REF. 3).

FT CONFLICT 99 99 R -> Q (IN REF. 3).

FT CONFLICT 270 270 M -> K (IN REF. 3).

FT CONFLICT 394 394 R -> Q (IN REF. 3).

FT CONFLICT 448 448 S -> G (IN REF. 2).

SQ SEQUENCE 576 AA; 62216 MW; 73EA8713DC54F85A CRC64;

Query Match 51.4%; Score 483.5; DB 1; Length 576;

Best Local Similarity 83.3%; Pred. No. 1.1e-35;

Matches 95; Conservative 0; Mismatches 10; Indels 9; Gaps 2;

QY 1 RGHWVGAHTLGHNSRGFVAIVGNVTAALPTEALRTVRDTPSCAVRAGLLRPDYALL 60

DB 461 RGHWVGAHTLGHNSRGFVAIVGNVTAALPTEALRTVRDTPSCAVRAGLLRPDYALL 520

QY 61 GHRQLVRTDCPGDALFLLRTWPHTAVSLRSLHYTARRP--SVYTSSTRPLPP 112

DB 521 GHRQLVRTDCPGDALFLLRTWPHTAT-----VKPRPARSVSKRSRREPPP 567

RESULT 2

PGPL PIG

ID PGPL PIG STANDARD; PRT; 598 AA.

AC Q866Y3; Q866Y4;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE N-acetylmuramoyl-L-alanine amidase precursor (EC 3.5.1.28)

DE (Peptidoglycan recognition protein long) (PGRP-L).

GN PGRP.L.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS A AND B).

RA Sang Y., Koss C.R., Blesha F.;

RT "Characterization of porcine peptidoglycan recognition proteins: gene

RT cloning and regulation on innate immunity."

RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: May plays a scavenger role by digesting biologically

CC active peptidoglycan (PGN) into biologically inactive fragments.

CC Has no direct bacteriolytic activity.

CC -!- CATALYTIC ACTIVITY: Hydrolyzes the link between N-acetylmuramoyl

CC residues and L-amino acid residues in certain bacterial cell-wall

CC glycopeptides.

CC -!- COFACTOR: Zinc (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted and membrane-associated.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=B;

CC IsoId=Q866Y3-1; Sequence=Displayed;

CC Name=A;

CC IsoId=Q866Y3-2; Sequence=VSP\_009082, VSP\_009083;

CC Note=No experimental confirmation available;

CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase

CC family 2.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF541955; AAO41115.1; -.

DR EMBL; AF541956; AAO41116.1; -.



DR InterPro; IPR002502; Amidase\_2.  
DR InterPro; IPR006619; PGRP.  
DR Pfam; PF01510; Amidase\_2; 1.  
DR SMART; SM00644; Ami\_2; 1.  
DR SMART; SM00701; PGRP; 1.  
KW Hydrolase; Immune response; Metal-binding; Zinc; Signal; Glycoprotein;  
KW Alternative splicing.  
FT SIGNAL 1 31  
FT CHAIN 32 598  
FT METAL 433 433  
FT METAL 469 469  
FT METAL 544 544  
FT METAL 552 552  
FT DISULFID 441 447  
FT CARBOHYD 353 353  
FT CARBOHYD 507 507  
FT VARSPLIC 1 346  
FT VARSPLIC 347 356  
FT SEQUENCE 598 AA; 64593 MW; FCDD237A9F105DDB CRC64;  
SQ  
Query Match 44.3%; Score 416.5; DB 1; Length 598;  
Best Local Similarity 73.0%; Pred. No. 1.1e-29;  
Matches 81; Conservative 8; Mismatches 17; Indels 5; Gaps 2;  
QY 1 RGHVWVGAHTLGHNSRGFGVAIVGNVTAALPTEAALTRVDTLPSCAVRAGLLRPDYALL 60  
Db 483 RGHVWVGAHTRDHNSRGFGVALIGNVTAELPSEALRAVRDELPHCAVRAGLLQPDYALL 542  
QY 61 GHRQLVTRDCPGDALFDLLRTWPHFTAVSLRSLHYTARRPSVYTSSTRPLP 111  
Db 543 GHRQLVTRDCPGDALFNMLRTWPRF---NMNVKPTARRAS--GRSKRRLP 588  
RESULT 3  
PGPL MOUSE STANDARD; PRT; 530 AA.  
ID Q8VCS0; Q8K4I8; Q9QXZ1; Q9QXZ2;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE N-acetylmuramoyl-L-alanine amidase precursor (EC 3.5.1.28)  
DE (peptidoglycan recognition protein long) (PGRP-L) (TagL).  
GN PGRPL OR PGLYRPL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 23-37.  
RC STRAIN=C57BL/6J;  
RX MEDLINE=22705304; PubMed=12821140;  
RA Gelius E., Persson C., Karlsson J., Steiner H.;  
RT "A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-L-alanine amidase activity."  
RL Biochem. Biophys. Res. Commun. 306:988-994 (2003).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE=22447357; PubMed=12559914;  
RA Kibardin A.V., Mirkina I.I., Baranova E.V., Zakeyeva I.R.,  
RA Georgiev G.P., Kiselev S.L.;  
RT "The differentially spliced mouse tagL gene, homolog of tag7/PGRP gene family in mammals and Drosophila, can recognize Gram-positive and Gram-negative bacterial cell wall independently of T phage lysozyme homology domain."  
RL J. Mol. Biol. 326:467-474 (2003).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Liver;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -!- FUNCTION: May plays a scavenger role by digesting biologically active peptidoglycan (PGN) into biologically inactive fragments.  
CC Has no direct bacteriolytic activity.  
CC -!- CATALYTIC ACTIVITY: Hydrolyzes the link between N-acetylmuramoyl residues and L-amino acid residues in certain bacterial cell-wall glycopeptides.  
CC -!- COFACTOR: Zinc (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted and membrane-associated.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=TagL-alpha;  
CC IsoId=Q8VCS0-1; Sequence=Displayed;  
CC Name=2; Synonyms=TagL-beta;  
CC IsoId=Q8VCS0-2; Sequence=VSP\_009081;  
CC Name=3; Synonyms=TagL-epsilon;  
CC IsoId=Q8VCS0-3; Sequence=VSP\_009079, VSP\_009080;  
CC -!- TISSUE SPECIFICITY: Strongly expressed in liver and fetal liver.  
CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase family 2.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AY282722; AAP22283.1; -  
EMBL; AF392055; AAM73674.1; -  
EMBL; AF149837; AAF22233.1; -  
EMBL; AF149838; AAF22234.1; -  
EMBL; BC019396; AAH19396.1; -  
MGD; MGI:1928099; Pglyrpl.  
InterPro; IPR002502; Amidase\_2.  
InterPro; IPR006619; PGRP.  
Pfam; PF01510; Amidase\_2; 1.  
SMART; SM00644; Ami\_2; 1.  
SMART; SM00701; PGRP; 1.  
KW Hydrolase; Immune response; Metal-binding; Zinc; Signal; Glycoprotein;  
KW Alternative splicing.  
FT SIGNAL 1 22  
FT CHAIN 23 530  
FT METAL 391 391  
FT METAL 427 427  
FT METAL 502 502  
FT METAL 510 510  
FT DISULFID 399 405  
FT CARBOHYD 61 61  
FT CARBOHYD 80 80  
FT CARBOHYD 174 174  
FT CARBOHYD 335 335  
FT CARBOHYD 465 465  
FT VARSPLIC 428 450  
N-ACETYLMURAMOYL-L-ALANINE AMIDASE.  
ZINC (BY SIMILARITY).  
ZINC (BY SIMILARITY).  
ZINC (BY SIMILARITY).  
ZINC (BY SIMILARITY).  
BY SIMILARITY.  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
SFVVGSDGYLYQGRGHWVGAHT -> RLTKNSFERPLKI  
QEVLSLML (in isoform 3).











OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 40-43; 50-55; 83-88; 127-145; 141-146  
RP AND 178-190, TISSUE DISTRIBUTION, FUNCTION, MASS SPECTROMETRY, AND  
RP PYRROLIDONE CARBOXYLIC ACID.  
RX MEDLINE=22028028; PubMed=11880375;  
RA Tydel C.C., Yount N., Tran D., Yuan J., Selsted M.E.;  
RT "Isolation, characterization, and antimicrobial properties of bovine  
RT oligosaccharide-binding protein. A microbicidal granule protein of  
RT eosinophils and neutrophils.";  
RL J. Biol. Chem. 277:19658-19664(2002).  
CC -!- FUNCTION: Involved in innate immunity. Is microbicidal for Gram-  
CC positive and Gram-negative bacteria and yeast.  
CC -!- SUBCELLULAR LOCATION: Secreted; cytoplasmic granules.  
CC -!- TISSUE SPECIFICITY: Synthesized only in bone marrow. The mature  
CC protein is stored in the cytoplasmic granules of eosinophils and  
CC neutrophils but is absent from monocytes, lymphocytes, or  
CC platelets.  
CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase  
CC family 2.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AY083309; AAL87002.1; -.  
CC InterPro; IPR002502; Amidase\_2.  
CC InterPro; IPR006619; PGRP.  
CC Pfam; PF01510; Amidase\_2; 1.  
CC SMART; SM00644; Ami\_2; 1.  
CC SMART; SM00701; PGRP; 1.  
KW Immune response; Antibiotic; Fungicide; Signal;  
KW Pyrrolidone carboxylic acid.  
FT SIGNAL 1 21 PEPTIDOGLYCAN RECOGNITION PROTEIN.  
FT CHAIN 22 190  
FT DISULFID 24 148  
FT DISULFID 40 85  
FT DISULFID 61 67  
FT MOD RES 22 22 PYRROLIDONE CARBOXYLIC ACID.  
FT SEQUENCE 190 AA; 21063 MW; 2BA7D659438F4ED7 CRC64;  
  
Query Match 18.3%; Score 172; DB 1; Length 190;  
Best Local Similarity 37.1%; Pred. No. 1.8e-08;  
Matches 33; Conservative 19; Mismatches 35; Indels 2; Gaps 2;  
  
QY 1 RGWHVVGAAHT-LGHNSRGFGVAIVGNVTAALPTAAALRTVDRDTLPSCAVRAGLLRPDYAL 59  
Db 103 RGWNTLGAHSGPTWNPIAIGISFMGNMHRVPPASALRAAQSL-ACGAARGYLTPNEYV 161  
  
QY 60 LGHRQLVTRDCPGDALFDLLRTWPHFTA 88  
Db 162 KGRDVGQTLSPGDLYKIIQQWPHYRV 190  
  
RESULT 10  
PGRP\_CAMDR  
ID\_PGRP\_CAMDR STANDARD; PRT; 193 AA.  
AC Q9GK12;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Peptidoglycan recognition protein precursor (Peptidoglycan recognition  
DE protein short) (PGRP-S).  
GN PGLYRP OR PGRP.  
OS Camelus dromedarius (Dromedary) (Arabian camel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
OX NCBI\_TaxID=9838;  
RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE=Lymphocytes;  
RA Kappeler S.R., Farah Z., Puhan Z.;  
RT "The peptidoglycan recognition protein is expressed in the lactating  
RT mammary gland of camels and binds to lactic acid bacteria.";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Binds specifically to peptidoglycan and is involved in  
CC innate immunity. Function in intracellular killing of bacteria  
CC (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted; cytoplasmic granules (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase  
CC family 2.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AJ131676; CAC19553.1; -.  
CC EMBL; AJ409286; CAC84130.1; -.  
CC InterPro; IPR002502; Amidase\_2.  
CC InterPro; IPR006619; PGRP.  
CC Pfam; PF01510; Amidase\_2; 1.  
CC SMART; SM00644; Ami\_2; 1.  
CC SMART; SM00701; PGRP; 1.  
KW Antibiotic; Immune response; Signal.  
FT SIGNAL 1 21 PEPTIDOGLYCAN RECOGNITION PROTEIN.  
FT CHAIN 22 193  
FT DISULFID 28 152 BY SIMILARITY.  
FT DISULFID 44 89 BY SIMILARITY.  
FT DISULFID 65 71 BY SIMILARITY.  
FT SEQUENCE 193 AA; 21377 MW; B6A1BD818030A7CB CRC64;  
  
Query Match 17.7%; Score 167; DB 1; Length 193;  
Best Local Similarity 38.6%; Pred. No. 5e-08;  
Matches 34; Conservative 17; Mismatches 35; Indels 2; Gaps 2;  
  
QY 1 RGWHVVGAAHT-LGHNSRGFGVAIVGNVTAALPTAAALRTVDRDTLPSCAVRAGLLRPDYAL 59  
Db 107 RGWNIKGAHAGPTWNEFISGISFMGNMNRVPPRALRAAQNL-ACGVALGALRSNEYV 165  
  
QY 60 LGHRQLVTRDCPGDALFDLLRTWPHFTA 87  
Db 166 KGRDVGQTLSPGDRLYEIIQTWSHYRA 193  
  
RESULT 11  
PGRP\_RAT  
ID\_PGRP\_RAT STANDARD; PRT; 183 AA.  
AC Q9JLN4;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Peptidoglycan recognition protein precursor (Peptidoglycan recognition  
DE protein short) (PGRP-S).  
GN PGLYRP OR PGRP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;  
RX MEDLINE=21066551; PubMed=11145837;  
RA Rehman A., Taishi P., Fang J., Majde J.A., Krueger J.M.;  
RT "The cloning of a rat peptidoglycan recognition protein (PGRP) and  
RT its induction in brain by sleep deprivation.";  
RL Cytokine 13:8-17(2001).  
CC -!- FUNCTION: Binds specifically to peptidoglycan and is involved in

```
CC      innate immunity. Function in intracellular killing of bacteria.
CC      -!- SUBCELLULAR LOCATION: Secreted; cytoplasmic granules (By
CC      similarity).
CC      -!- TISSUE SPECIFICITY: Expressed in all regions of the brain.
CC      -!- INDUCTION: By sleep deprivation in the brain stem and in the
CC      hypothalamus.
CC      -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase
CC      family 2.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AF154114; AAF73252.1; -.
CC      InterPro; IPR002502; Amidase_2.
CC      InterPro; IPR006619; PGRP.
CC      Pfam; PF01510; Amidase_2; 1.
CC      SMART; SM00644; Ami_2; 1.
CC      SMART; SM00701; PGRP; 1.
CC      Antibiatic; Immune response; Signal.
CC      SIGNAL 1 17 POTENTIAL.
CC      CHAIN 18 183 PEPTIDOGLYCAN RECOGNITION PROTEIN.
CC      DISULFID 18 142 BY SIMILARITY.
CC      FT DISULFID 34 79 BY SIMILARITY.
CC      FT DISULFID 55 61 BY SIMILARITY.
CC      SQ SEQUENCE 183 AA; 20590 MW; 5B9C1B7AA8A2EC21 CRC64;
CC
CC      Query Match 16.5%; Score 155; DB 1; Length 183;
CC      Best Local Similarity 36.8%; Pred. No. 5.5e-07;
CC      Matches 32; Conservative 17; Mismatches 34; Indels 4; Gaps 3;
CC
Qy      1 RGHWVGAHTLG--HNSRGFGVAIVGNVTAALPTEAALRTVRDITLPSCAVRAGLLRPDYA 58
Db      97 RGWTTKGDHT-GPIWNPMISIGITFWGDSHRVPAKRALRAALNLL-KCGVSEGFRLRSNYE 154
Qy      59 LLGHRQLVRTDCPGDALFDLLRTWPHF 85
Db      155 VKGHRDVQSTLSPGDQLYEIIQSWDHY 181
RESULT 12
PGRP_MOUSE STANDARD; PRT; 182 AA.
AC      O88593; Q62185;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Peptidoglycan recognition protein precursor (Peptidoglycan recognition
DE      protein short) (PGRP-S) (Cytokine tag7).
GN      PGLYRP OR PGRP OR PGRPS OR TAG7.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC      TISSUE=Spleen;
RX      MEDLINE=98374308; PubMed=9707603;
RA      Kang D.; Liu G.; Lundstroem A.; Gelius E.; Steiner H.;
RT      "A peptidoglycan recognition protein in innate immunity conserved from
RT      insects to humans.";
RL      Proc. Natl. Acad. Sci. U.S.A. 95:10078-10082(1998).
RN      [2]
RP      SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
RX      MEDLINE=98325081; PubMed=9660837;
RA      Kiselev S.L.; Kustikova O.S.; Korobko E.V.; Prokhortchouk E.B.;
RA      Kabishev A.A.; Lukanidin E.M.; Georgiev G.P.;
RT      "Molecular cloning and characterization of the mouse tag7 gene
RT      encoding a novel cytokine.";
```

```
RL      J. Biol. Chem. 273:18633-18639(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Slayton W.B.; Rigaa A.; Hancock J.D.; Zaugg J.K.; Le T.V.;
RA      Trautman M.S.; Spangrude G.J.; Carroll W.L.; Schibler K.R.;
RT      "Granulocyte-colony stimulating factor up-regulates expression of
RT      murine tag7 during myeloid differentiation.";
RL      Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A., AND FUNCTION.
RC      STRAIN=129/Sv;
RX      MEDLINE=22718676; PubMed=12649138;
RA      Dziarski R.; Platt K.A.; Gelius E.; Steiner H.; Gupta D.;
RT      "Defect in neutrophil killing and increased susceptibility to
RT      infection with nonpathogenic Gram-positive bacteria in peptidoglycan
RT      recognition protein-S (PGRP-S)-deficient mice.";
RL      Blood 102:689-697(2003).
RN      [5]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Small intestine;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J.; Shinagawa A.; Shibata K.; Yoshino M.; Itoh M.; Ishii Y.;
RA      Arakawa T.; Hara A.; Fukunishi Y.; Konno H.; Adachi J.; Fukuda S.;
RA      Aizawa K.; Izawa M.; Nishi K.; Kiyosawa H.; Kondo S.; Yamanaka I.;
RA      Saito T.; Okazaki Y.; Gojobori T.; Bono H.; Kasukawa T.; Saito R.;
RA      Kadota K.; Matsuda H.A.; Ashburner M.; Batalov S.; Casavant T.;
RA      Fleischmann W.; Gaasterland T.; Gissi C.; King B.; Kochiwa H.;
RA      Kuehl P.; Lewis S.; Matsuo Y.; Nikaido I.; Pesole G.; Quackenbush J.;
RA      Schriml L.M.; Staubli F.; Suzuki R.; Tomita M.; Wagner L.; Washio T.;
RA      Sakai K.; Okido T.; Furuno M.; Aono H.; Baldarelli R.; Barsh G.;
RA      Blake J.; Boffelli D.; Bojunga N.; Carninci P.; de Bonaldo M.F.;
RA      Brownstein M.J.; Bult C.; Fletcher C.; Fujita M.; Gariboldi M.;
RA      Gustincich S.; Hill D.; Hofmann M.; Hume D.A.; Kamiya M.; Lee N.H.;
RA      Lyons P.; Marchionni L.; Mashima J.; Mazzarelli J.; Mombaerts P.;
RA      Nordone P.; Ring B.; Ringwald M.; Rodriguez I.; Sakamoto N.;
RA      Sasaki H.; Sato K.; Schoenbach C.; Seya T.; Shibata Y.; Storch K.-F.;
RA      Suzuki H.; Toyooka K.; Wang K.H.; Weitz C.; Whittaker C.; Wilming L.;
RA      Wynshaw-Boris A.; Yoshida K.; Hasegawa Y.; Kawaji H.; Kohtsuki S.;
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
RN      [6]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA      Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA      Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA      Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA      Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA      Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA      Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA      Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;
RA      Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA      Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA      Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA      Fahey J.; Helton E.; Kettman M.; Madan A.; Rodrigues S.; Sanchez A.;
RA      Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA      Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA      Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.;
RA      Butterfield Y.S.N.; Krzywinski M.I.; Skalska U.; Smailus D.E.;
RA      Schnerch A.; Schein J.E.; Jones S.J.M.; Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC      -!- FUNCTION: Binds specifically to peptidoglycan and is involved in
CC      innate immunity. Function in intracellular killing of bacteria.
CC      The soluble form triggers apoptosis in vitro.
CC      -!- SUBCELLULAR LOCATION: Exists in both soluble and membrane-
CC      associated forms.
CC      -!- TISSUE SPECIFICITY: Strongly expressed in spleen and lung. Also
CC      detected in brain and thymus. In the lung, expressed in the
CC      intraalveolar space, in the brain, expressed in the Purkinje cells
CC      of the cerebellum and in certain layers of neurons in the
```





```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M82323; AAB62389.1; -.
DR EMBL; M80627; AAA58632.1; -.
DR EMBL; M65209; AAC37571.1; -.
DR PIR; A42121; A42121.
DR TRANSFAC; T01503; -.
DR Genew; HGNC:11623; TCF12.
DR MIM; 600480; -.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0007517; P:muscle development; TAS.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; TAS.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Developmental protein.
FT DOMAIN 119 140 LEUCINE-ZIPPER (POTENTIAL).
FT DNA_BIND 577 589 BASIC DOMAIN.
FT DOMAIN 590 631 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 632 655 CLASS A SPECIFIC DOMAIN.
SQ SEQUENCE 682 AA; 72964 MW; 9736113D9361D3F5 CRC64;

Query Match 8.9%; Score 84; DB 1; Length 682;
Best Local Similarity 28.2%; Pred. No. 5;
Matches 35; Conservative 13; Mismatches 52; Indels 24; Gaps 6;

QY 40 RDT-LPSCAVRAGLLRPDYALLGHRQLVTRDCPGDALFDLLRTWPHFTAVS--LRSLHYT 96
Db 122 RDTGLPGC--QSSLLRQLGLGSPAQLSSSGKPGTAYYS-----FSATSSRRRPLHDS 172

QY 97 ARRPVSVYTSSTRPLPACNSCARTASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPV TSA 156
Db 173 AALDPLQAKKVKVPGLPSSVYAPSPNSDDFNRE-----SPSYPS-----PKPPTSM 220

QY 157 YAAS 160
Db 221 FAST 224

RESULT 15
FXD1_HUMAN STANDARD; PRT; 465 AA.
AC Q16676; Q12949;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Forkhead box protein D1 (Forkhead-related protein FKHL8) (Forkhead-
DE related transcription factor 4) (FREAC-4).
GN FOXD1 OR FKHL8 OR FREAC4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355467; PubMed=8702877;
RA Ernstsson S., Pierrou S., Hulander M., Cederberg A., Hellqvist M.,
RA Carlsson P., Enerback S.;
RT "Characterization of the human forkhead gene FREAC-4. Evidence for
RT regulation by Wilms' tumor suppressor gene (WT-1) and p53.";
RL J. Biol. Chem. 271:21094-21099(1996).
RN [2]
RP SEQUENCE OF 120-225 FROM N.A.
RX MEDLINE=95045392; PubMed=7957066;
RA Pierrou S., Hellqvist M., Samuelsson L., Enerbaeck S., Carlsson P.;
RA "Cloning and characterization of seven human forkhead proteins:
RT binding site specificity and DNA bending.";
RL EMBO J. 13:5002-5012(1994).
```

```
CC -!- FUNCTION: Binding of FREAC-3 and FREAC-4 to their cognate sites
CC results in bending of the DNA at an angle of 80-90 degrees.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 fork-head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U59832; AAC50661.1; -.
DR EMBL; U59831; AAC50660.1; -.
DR EMBL; U13222; AAA92039.1; -.
DR PIR; G02738; G02738.
DR PIR; S51627; S51627.
DR HSSP; Q63245; 2HEH.
DR TRANSFAC; T02472; -.
DR Genew; HGNC:3802; FOXD1.
DR MIM; 601091; -.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS50039; FORK_HEAD_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DOMAIN 26 34 POLY-GLU.
FT DOMAIN 39 43 POLY-GLY.
FT DOMAIN 52 57 POLY-ARG.
FT DOMAIN 69 72 POLY-GLU.
FT DOMAIN 73 76 POLY-ASP.
FT DOMAIN 97 113 POLY-GLY.
FT DNA_BIND 124 215 FORK-HEAD.
FT DOMAIN 231 234 POLY-ALA.
FT DOMAIN 252 256 POLY-ALA.
FT DOMAIN 259 266 POLY-PRO.
FT DOMAIN 293 303 POLY-ALA.
FT DOMAIN 309 315 POLY-PRO.
FT DOMAIN 375 390 POLY-ALA.
FT DOMAIN 428 434 POLY-ALA.
SQ SEQUENCE 465 AA; 46140 MW; D3E7854909CCBF AE CRC64;

Query Match 8.9%; Score 83.5; DB 1; Length 465;
Best Local Similarity 30.9%; Pred. No. 3.6;
Matches 30; Conservative 9; Mismatches 47; Indels 11; Gaps 3;

QY 83 PHFTAVSLRSLHYTARRPSVYTSSTRPLPACNSC---ARTASARPPTSRHHVYSGNLGP 139
Db 315 PHGAAAEELARTAFGYRPHPLGAALPGPLPASAAKAGPGCASALARSPPFSIESIIGSLGP 374

QY 140 AFAGHSAGNIPDPVTSAYAAQAQPTQP--ACPFPPSS 174
Db 375 ANAAAAAAQ-----AAAAAQASPPSPVAAAPPAGS 405

Search completed: May 18, 2004, 16:12:02
Job time : 19 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:13:34 ; Search time 59 Seconds  
(without alignments)  
833.276 Million cell updates/sec

Title: US\_10-068-956-2  
Perfect score: 174  
Sequence: 1 RGHWVGAHTLGHNSRGFV.....SAYAASAQPTQPACFPFPSS 174

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	174	100.0	174	AA72663	Aay72663 Human pep
2	174	100.0	634	AAG65915	Aag65915 Amino aci
3	174	100.0	634	ADA54695	Ada54695 Human pro
4	87	50.0	363	AA94863	Aay94863 Human pro
5	87	50.0	576	AAG65916	Aag65916 Amino aci
6	87	50.0	576	ABB78298	Abb78298 Amino aci
7	87	50.0	576	ABR57566	Abr57566 Human Pep
8	87	50.0	576	ADE07869	Ade07869 Novel pro
9	10	5.7	530	AA72664	Aay72664 Murine pe
10	8	4.6	107	AAU40317	Aau40317 Propionib
11	8	4.6	107	ABM36836	Abm36836 Propionib
12	8	4.6	126	ABG17167	Abg17167 Novel hum
13	8	4.6	173	ABB63392	Abb63392 Drosophil
14	8	4.6	179	AAG35998	Aag35998 Zea mays
15	8	4.6	526	AA700051	Aay00051 Enterococ
16	8	4.6	526	ABP43270	Abp43270 E faecali
17	8	4.6	526	ABU13549	Abu13549 Enterococ
18	8	4.6	526	ABU88298	Abu88298 E. faecal
19	8	4.6	546	AA700050	Aay00050 Enterococ
20	8	4.6	546	ABP43269	Abp43269 E faecali
21	8	4.6	546	ABU88297	Abu88297 E. faecal
22	8	4.6	546	ABU13548	Abu13548 Enterococ
23	8	4.6	1197	3 AAY57445	Aay57445 Mouse Ese
24	8	4.6	1658	3 AAY57450	Aay57450 Mouse Ese
25	8	4.6	1683	3 AAY71160	Aay71160 Rat phosp

99	7	4.0	554	7	ADB91810	Adb91810	Human sec	172	6	3.4	43	6	ABU13134	Abu13134	Novel hum
100	7	4.0	554	7	ADC74556	Adc74556	Human sec	173	6	3.4	46	3	AAB34636	Aab34636	Human sec
101	7	4.0	561	3	AAG40282	Aag40282	Arabidops	174	6	3.4	46	3	AAG58760	Aag58760	Arabidops
102	7	4.0	594	4	AAM933619	Aam933619	Human pol	175	6	3.4	50	2	AAW16936	Aaw16936	Random re
103	7	4.0	669	5	ABB92210	Abb92210	Herbicida	176	6	3.4	50	2	AAW25499	Aaw25499	Random pe
104	7	4.0	741	4	ABB60912	Abb60912	Drosophil	177	6	3.4	50	4	AAU40043	Aau40043	Propionib
105	7	4.0	783	2	AAAY60344	Aay60344	Human nor	178	6	3.4	50	4	AAU61413	Aau61413	Propionib
106	7	4.0	791	4	ABG12750	Abg12750	Novel hum	179	6	3.4	50	6	ABM36562	Abm36562	Propionib
107	7	4.0	838	4	ABB70548	Abb70548	Drosophil	180	6	3.4	50	6	ABM57932	Abm57932	Propionib
108	7	4.0	853	4	ABG12054	Abg12054	Novel hum	181	6	3.4	51	5	ABP08811	Abp08811	Human ORF
109	7	4.0	889	4	ABB58940	Abb58940	Drosophil	182	6	3.4	51	6	ABP81017	Abp81017	N. gonorr
110	7	4.0	929	6	AAE36026	Aae36026	Human mem	183	6	3.4	52	3	AAAY65391	Aay65391	Human 5'
111	7	4.0	950	2	AAAY33298	Aay33298	Human mem	184	6	3.4	52	4	AAU65449	Aau65449	Propionib
112	7	4.0	950	5	ABB81194	Abb81194	Human mem	185	6	3.4	52	4	AAU49296	Aau49296	Propionib
113	7	4.0	950	6	ADA57193	Ada57193	Human sec	186	6	3.4	52	4	AAU64194	Aau64194	Propionib
114	7	4.0	950	6	AAE36025	Aae36025	Human mem	187	6	3.4	52	6	ABM45815	Abm45815	Propionib
115	7	4.0	950	6	ADA41070	Ada41070	Human sec	188	6	3.4	52	6	ABM60713	Abm60713	Propionib
116	7	4.0	950	6	ABR47911	Abrr47911	Human sec	189	6	3.4	52	6	ABM61968	Abm61968	Propionib
117	7	4.0	950	6	ABR00162	Abrr00162	Human gen	190	6	3.4	53	3	AAG58759	Aag58759	Arabidops
118	7	4.0	950	7	ADB91654	Adb91654	Human sec	191	6	3.4	53	4	AAU61178	Aau61178	Propionib
119	7	4.0	950	7	ADC74310	Adc74310	Human sec	192	6	3.4	53	6	ABM57697	Abm57697	Propionib
120	7	4.0	950	7	ADC78245	Adc78245	Human sec	193	6	3.4	53	7	ADC03398	Adc03398	Rice flow
121	7	4.0	1008	4	AAAB82247	Aab82247	Rat insul	194	6	3.4	54	4	AAU22110	Aau22110	Human car
122	7	4.0	1008	7	ABW01163	Abw01163	Human IRD	195	6	3.4	54	7	ADE46078	Ade46078	Human car
123	7	4.0	1008	7	ABW01138	Abw01138	Rat IRDBP	196	6	3.4	56	4	AAM84072	Aam84072	Human imm
124	7	4.0	1028	4	AAAB82249	Aab82249	Human ins	197	6	3.4	56	4	AAU67515	Aau67515	Propionib
125	7	4.0	1047	7	ABW01140	Abw01140	Alternati	198	6	3.4	56	4	AAU51856	Aau51856	Propionib
126	7	4.0	1054	7	ABW01144	Abw01144	Rat IRDBP	199	6	3.4	56	4	AAU40061	Aau40061	Propionib
127	7	4.0	1186	6	ABR53096	Abrr53096	Protein s	200	6	3.4	56	4	AAU49926	Aau49926	Propionib
128	7	4.0	1288	7	ADE16050	Adel16050	G-coupled	201	6	3.4	56	6	ABM46445	Abm46445	Propionib
129	7	4.0	1408	7	ADE16052	Adel16052	G-coupled	202	6	3.4	56	6	ABM36580	Abm36580	Propionib
130	7	4.0	1463	7	ABW01145	Abw01145	Human IRD	203	6	3.4	56	6	ABM64034	Abm64034	Propionib
131	7	4.0	1516	4	AAAB11435	Aab11435	C. albica	204	6	3.4	56	6	ABM48375	Abm48375	Propionib
132	7	4.0	2616	4	ABB64889	Abb64889	Drosophil	205	6	3.4	58	4	AAU61410	Aau61410	Propionib
133	6	3.4	7	4	AAAM47017	Aam47017	H11 bindi	206	6	3.4	58	6	ABM57929	Abm57929	Propionib
134	6	3.4	7	4	AAAM46962	Aam46962	H11 bindi	207	6	3.4	59	7	ADC03422	Adc03422	Rice flow
135	6	3.4	7	4	AAAM46957	Aam46957	H11 bindi	208	6	3.4	61	3	AAB09341	Aab09341	Hepatitis
136	6	3.4	7	4	AAAM47075	Aam47075	H11 bindi	209	6	3.4	61	4	AAU66438	Aau66438	Propionib
137	6	3.4	7	4	AAAM46952	Aam46952	H11 bindi	210	6	3.4	61	6	ABM62957	Abm62957	Propionib
138	6	3.4	10	2	AAAR93333	Aar93333	YES prote	211	6	3.4	62	3	AAB33404	Aab33404	Pinus rad
139	6	3.4	11	6	AAAM00580	Aam00580	Human tra	212	6	3.4	62	4	AAU57631	Aau57631	Propionib
140	6	3.4	11	6	ABR39758	Abrr39758	Subtilisi	213	6	3.4	62	6	ABM54150	Abm54150	Propionib
141	6	3.4	12	4	AAAB50778	Aab50778	Human cAM	214	6	3.4	63	4	AAU67882	Aau67882	Propionib
142	6	3.4	13	2	AAAW11107	Aaw11107	Src SH3 d	215	6	3.4	63	6	ABM64401	Abm64401	Propionib
143	6	3.4	13	2	AAAW11123	Aaw11123	Src SH3 d	216	6	3.4	64	4	ABB40758	Abb40758	Peptide #
144	6	3.4	13	2	AAAW11109	Aaw11109	Src SH3 d	217	6	3.4	64	4	AAAM34524	Aam34524	Peptide #
145	6	3.4	13	2	AAAW11104	Aaw11104	Src SH3 d	218	6	3.4	64	4	AAAM74410	Aam74410	Human bon
146	6	3.4	14	2	AAAR93425	Aar93425	GST-YES S	219	6	3.4	64	4	AAAM61618	Aam61618	Human bra
147	6	3.4	14	2	AAAR933454	Aar93454	GST-PI3K	220	6	3.4	64	4	ABG56205	Abg56205	Human liv
148	6	3.4	14	2	AAAR93384	Aar93384	GST-SRC S	221	6	3.4	64	5	ABG44312	Abg44312	Human pep
149	6	3.4	14	2	AAAW88650	Aaw88650	Secreted	222	6	3.4	65	4	AAO12080	Aao12080	Human pol
150	6	3.4	14	4	ABB50417	Abb50417	Human sec	223	6	3.4	65	4	AAU40839	Aau40839	Propionib
151	6	3.4	14	6	ABO44674	Abo44674	Novel hum	224	6	3.4	65	6	ABM37358	Abm37358	Propionib
152	6	3.4	14	7	ABO26154	Abo26154	Human pro	225	6	3.4	68	4	AAU65388	Aau65388	Propionib
153	6	3.4	15	2	AAAW38993	Aaw38993	Peptide r	226	6	3.4	68	4	AAU45849	Aau45849	Propionib
154	6	3.4	16	2	AAAW39012	Aaw39012	Peptide r	227	6	3.4	68	6	ABM42368	Abm42368	Propionib
155	6	3.4	17	2	AAAW38973	Aaw38973	Peptide r	228	6	3.4	68	6	ABM61907	Abm61907	Propionib
156	6	3.4	18	6	ADA26923	Ada26923	TRANSFACM	229	6	3.4	69	2	AAW05553	Aaw05553	Coxsackie
157	6	3.4	24	4	AAAB87099	Aab87099	Human TAN	230	6	3.4	69	2	AAW05552	Aaw05552	Coxsackie
158	6	3.4	25	7	ADD90506	Add90506	Novel hum	231	6	3.4	69	2	AAW05545	Aaw05545	Coxsackie
159	6	3.4	30	6	ABR39745	Abrr39745	ADNF I po	232	6	3.4	69	3	AAG18235	Aag18235	Arabidops
160	6	3.4	31	2	AAAW16939	Aaw16939	Random re	233	6	3.4	69	4	AAU66181	Aau66181	Propionib
161	6	3.4	31	2	AAAW16942	Aaw16942	Random re	234	6	3.4	69	6	ABM62700	Abm62700	Propionib
162	6	3.4	31	2	AAAW25505	Aaw25505	Random pe	235	6	3.4	70	3	AAG56022	Aag56022	Arabidops
163	6	3.4	31	2	AAAW25502	Aaw25502	Random pe	236	6	3.4	70	4	AAU46265	Aau46265	Propionib
164	6	3.4	32	6	ABP80231	Abp80231	N. gonorr	237	6	3.4	70	6	ABP80197	Abp80197	N. gonorr
165	6	3.4	35	5	ABP28346	Abp28346	Streptoco	238	6	3.4	70	6	ABM42784	Abm42784	Propionib
166	6	3.4	41	4	AAAM34283	Aam34283	Peptide #	239	6	3.4	71	4	AAAM88156	Aam88156	Human imm
167	6	3.4	41	4	AAAM74136	Aam74136	Human bon	240	6	3.4	72	2	AAW42444	Aaw42444	S. commun
168	6	3.4	41	4	AAAM61379	Aam61379	Human bra	241	6	3.4	72	4	AAU45631	Aau45631	Propionib
169	6	3.4	41	4	ABG55911	Abg55911	Human liv	242	6	3.4	72	6	ABM42150	Abm42150	Propionib
170	6	3.4	41	5	ABG44053	Abg44053	Human pep	243	6	3.4	72	7	ADC35027	Adc35027	Ras-like
171	6	3.4	43	4	ABB03840	Abb03840	Human mus	244	6	3.4	73	2	AAAR53253	Aar53253	Signal pe

172	6	3.4	43	6	ABU13134	Abu13134	Novel hum
173	6	3.4	46	3	AAB34636	Aab34636	Human sec
174	6	3.4	46	3	AAG58760	Aag58760	Arabidops
175	6	3.4	50	2	AAW16936	Aaw16936	Random re
176	6	3.4	50	2	AAW25499	Aaw25499	Random pe
177	6	3.4	50	4	AAU40043	Aau40043	Propionib
178	6	3.4	50	4	AAU61413	Aau61413	Propionib
179	6	3.4	50	6	ABM36562	Abm36562	Propionib
180	6	3.4	50	6	ABM57932	Abm57932	Propionib
181	6	3.4	51	5	ABP08811	Abp08811	Human ORF
182	6	3.4	51	6	ABP81017	Abp81017	N. gonorr
183	6	3.4	52	3	AAAY65391	Aay65391	Human 5'
184	6	3.4	52	4	AAU65449	Aau65449	Propionib
185	6	3.4	52	4	AAU49296	Aau49296	Propionib
186	6	3.4	52	4	AAU64194	Aau64194	Propionib
187	6	3.4	52	6	ABM45815	Abm45815	Propionib
188	6	3.4	52	6	ABM60713	Abm60713	Propionib
189	6	3.4	52	6	ABM61968	Abm61968	Propionib
190	6	3.4	53	3	AAG58759	Aag58759	Arabidops
191	6	3.4	53	4	AAU61178	Aau61178	Propionib
192	6	3.4	53	6	ABM57697	Abm57697	Propionib
193	6	3.4	53	7	ADC03398	Adc03398	Rice flow
194	6	3.4	54	4	AAU22110	Aau22110	Human car
195	6	3.4	54	7	ADE46078	Ade46078	Human car
196	6	3.4	56	4	AAM84072	Aam84072	Human imm
197	6	3.4	56	4	AAU67515	Aau67515	Propionib
198	6	3.4	56	4	AAU51856	Aau51856	Propionib
199	6	3.4	56	4	AAU40061	Aau40061	Propionib
200	6	3.4	56	4	AAU49926	Aau49926	Propionib
201	6	3.4	56	6	ABM46445	Abm46445	Propionib
202	6	3.4	56	6	ABM36580	Abm36580	Propionib
203	6	3.4	56	6	ABM64034	Abm64034	Propionib



245 6 3.4 73 3 AAY76985 Aay76985 Corn star 318 6 3.4 97 4 ABG24258 Abg24258 Novel hum  
246 6 3.4 74 4 AAU65172 Aau65172 Propionib 319 6 3.4 97 6 ABU06364 Abu06364 Maize Sta  
247 6 3.4 74 6 ABM61691 Abm61691 Propionib 320 6 3.4 97 7 ADD28122 Add28122 Lymphoma  
248 6 3.4 76 4 AAU56846 Aau56846 Propionib 321 6 3.4 98 2 AAU99666 Aaw99666 Human sec  
249 6 3.4 76 6 ABM53365 Abm53365 Propionib 322 6 3.4 98 4 AAO02066 Aao02066 Human pol  
250 6 3.4 77 4 AAG77638 Aag77638 Human col 323 6 3.4 98 4 ABB16157 Abb16157 Human ner  
251 6 3.4 77 4 AAU59885 Aau59885 Propionib 324 6 3.4 98 5 ABP05242 Abp05242 Human ORF  
252 6 3.4 77 6 ABM56404 Abm56404 Propionib 325 6 3.4 98 6 ABU69154 Abu69154 Human NOV  
253 6 3.4 77 7 ADC32806 Adc32806 Human nov 326 6 3.4 98 7 ADA45101 Ada45101 Human pol  
254 6 3.4 79 3 AAG26097 Aag26097 Zea mays 327 6 3.4 99 3 AAG04125 Aag04125 Arabidops  
255 6 3.4 80 4 AAU64013 Aau64013 Propionib 328 6 3.4 99 3 AAG38378 Aag38378 Arabidops  
256 6 3.4 80 5 ABP35287 Abp35287 Human onc 329 6 3.4 101 4 AAU61535 Aau61535  
257 6 3.4 80 6 ABM60532 Abm60532 Propionib 330 6 3.4 101 5 ABP04030 Abp04030 Human ORF  
258 6 3.4 80 7 ADC94184 Adc94184 E. faeciu 331 6 3.4 101 6 ABM58054 Abm58054 Propionib  
259 6 3.4 80 7 AAM92577 Aam92577 Human dig 332 6 3.4 102 4 AAU68007 Aau68007 Propionib  
260 6 3.4 82 4 AAU62782 Aau62782 Propionib 333 6 3.4 102 5 AAO17456 Aao17456 Human liv  
261 6 3.4 82 4 AAU46825 Aau46825 Propionib 334 6 3.4 102 6 ABM64526 Abm64526 Propionib  
262 6 3.4 82 4 AAU62782 Aau62782 Propionib 335 6 3.4 102 7 ADC31594 Adc31594 Human nov  
263 6 3.4 82 4 AAU58216 Aau58216 Propionib 336 6 3.4 103 5 ABG60203 Abg60203 Human DIT  
264 6 3.4 82 4 AAU22603 Aau22603 Novel hum 337 6 3.4 104 3 AAG18233 Aag18233 Arabidops  
265 6 3.4 82 5 ABP63864 Abp63864 Human ORF 338 6 3.4 104 6 ABU06365 Abu06365 Maize Sta  
266 6 3.4 82 6 ABM59301 Abm59301 Propionib 339 6 3.4 105 3 AAB40258 Aab40258 Human ORF  
267 6 3.4 82 6 ABM43344 Abm43344 Propionib 340 6 3.4 105 3 AAG56021 Aag56021 Arabidops  
268 6 3.4 82 6 ABM54735 Abm54735 Propionib 341 6 3.4 105 4 AAU66611 Aau66611 Propionib  
269 6 3.4 82 7 ADB32443 Adb32443 Human nov 342 6 3.4 105 6 ABM63130 Abm63130 Propionib  
270 6 3.4 82 7 ADC35028 Adc35028 Ras-like 343 6 3.4 105 7 ADD29464 Add29464 Mouse cae  
271 6 3.4 83 5 AAU87817 Aau87817 Jojoba KC 344 6 3.4 107 3 AAG32882 Aag32882 Zea mays  
272 6 3.4 84 4 AAU58914 Aau58914 Propionib 345 6 3.4 107 3 AAG59688 Aag59688 Zea mays  
273 6 3.4 84 4 AAB65013 Aab65013 Human sec 346 6 3.4 111 4 AAU65503 Aau65503 Propionib  
274 6 3.4 84 6 ABM55433 Abm55433 Propionib 347 6 3.4 111 4 AAU31264 Aau31264 Novel hum  
275 6 3.4 85 3 AAB59062 Aab59062 Breast an 348 6 3.4 111 6 ABM62022 Abm62022 Propionib  
276 6 3.4 85 3 AAG27127 Aag27127 Zea mays 349 6 3.4 112 3 AAB40545 Aab40545 Human ORF  
277 6 3.4 85 3 AAG18234 Aag18234 Arabidops 350 6 3.4 112 4 AAM24343 Aam24343 Human EST  
278 6 3.4 85 4 AAO11066 Aao11066 Human pol 351 6 3.4 112 5 ABP01461 Abp01461 Human ORF  
279 6 3.4 86 4 AAU50886 Aau50886 Propionib 352 6 3.4 113 2 AAY73957 Aay73957 Human pro  
280 6 3.4 86 4 AAU62438 Aau62438 Propionib 353 6 3.4 113 4 AAU40352 Aau40352 Propionib  
281 6 3.4 86 5 ABP64976 Abp64976 Human pro 354 6 3.4 113 4 AAU30876 Aau30876 Novel hum  
282 6 3.4 86 6 ABM47405 Abm47405 Propionib 355 6 3.4 113 6 ABM36871 Abm36871 Propionib  
283 6 3.4 86 6 ABM58957 Abm58957 Propionib 356 6 3.4 115 3 AAG04124 Aag04124 Arabidops  
284 6 3.4 87 3 AAG00354 Aag00354 Human sec 357 6 3.4 115 3 AAG38377 Aag38377 Arabidops  
285 6 3.4 87 4 AAU56391 Aau56391 Propionib 358 6 3.4 116 2 AAR25593 Aar25593 Mutant RI  
286 6 3.4 87 4 AAU55826 Aau55826 Propionib 359 6 3.4 116 4 ABG10851 Abg10851 Novel hum  
287 6 3.4 87 6 ABM52345 Abm52345 Propionib 360 6 3.4 116 6 ABP76689 Abp76689 Streptomy  
288 6 3.4 87 6 ABM52910 Abm52910 Propionib 361 6 3.4 119 4 AAM24301 Aam24301 Human EST  
289 6 3.4 88 5 ABP34557 Abp34557 Human ORF 362 6 3.4 119 4 AAU63528 Aau63528 Propionib  
290 6 3.4 89 5 ABP00952 Abp00952 Human ORF 363 6 3.4 119 4 ABG23503 Abg23503 Novel hum  
291 6 3.4 90 4 AAM92567 Aam92567 Human dig 364 6 3.4 119 6 ABM60047 Abm60047 Propionib  
292 6 3.4 90 4 AAU22593 Aau22593 Novel hum 365 6 3.4 120 2 AAW74835 Aaw74835 Human sec  
293 6 3.4 90 7 ADB32433 Adb32433 Human nov 366 6 3.4 120 5 ABG95286 Abg95286 Human nov  
294 6 3.4 91 3 AAG54519 Aag54519 Zea mays 367 6 3.4 120 6 ABO34480 Abo34480 Region of  
295 6 3.4 91 4 AAU31136 Aau31136 Novel hum 368 6 3.4 121 4 AAB94940 Aab94940 Human pro  
296 6 3.4 92 4 AAG76407 Aag76407 Human col 369 6 3.4 121 4 AAG91805 Aag91805 C glutami  
297 6 3.4 92 4 AAU43032 Aau43032 Propionib 370 6 3.4 122 4 AAU51169 Aau51169 Propionib  
298 6 3.4 92 4 AAU40100 Aau40100 Propionib 371 6 3.4 122 5 ABP55524 Abp55524 Human lip  
299 6 3.4 92 4 AAU66244 Aau66244 Propionib 372 6 3.4 122 6 ABM47688 Abm47688 Propionib  
300 6 3.4 92 5 ABP35079 Abp35079 Human syn 373 6 3.4 122 6 ABM65163 Abm65163 Propionib  
301 6 3.4 92 6 ABM62763 Abm62763 Propionib 374 6 3.4 123 4 AAU41835 Aau41835 Propionib  
302 6 3.4 92 6 ABM36619 Abm36619 Propionib 375 6 3.4 123 5 ABP09027 Abp09027 Human ORF  
303 6 3.4 92 6 ABM39551 Abm39551 Propionib 376 6 3.4 123 6 ABM38354 Abm38354 Propionib  
304 6 3.4 93 3 AAB56956 Aab56956 Human pro 377 6 3.4 123 7 ADC94531 Adc94531 E. faeciu  
305 6 3.4 93 4 AAU40675 Aau40675 Propionib 378 6 3.4 124 3 AAG37466 Aag37466 Arabidops  
306 6 3.4 93 6 ABM37194 Abm37194 Propionib 379 6 3.4 124 5 ABP64087 Abp64087 Human ORF  
307 6 3.4 93 6 ABP75803 Abp75803 Human sec 380 6 3.4 127 3 AAG12633 Aag12633 Arabidops  
308 6 3.4 93 7 ADD90344 Add90344 Novel hum 381 6 3.4 127 3 AAG13266 Aag13266 Arabidops  
309 6 3.4 94 4 AAG74950 Aag74950 Human col 382 6 3.4 127 3 AAG34861 Aag34861 Arabidops  
310 6 3.4 95 2 AAY76563 Aay76563 Human ova 383 6 3.4 128 4 ABB17976 Abb17976 Human ner  
311 6 3.4 95 4 ABG23707 Abg23707 Novel hum 384 6 3.4 130 4 AAU19326 Aau19326 Human G p  
312 6 3.4 95 5 ABP01494 Abp01494 Human ORF 385 6 3.4 130 4 AAM88142 Aam88142 Human imm  
313 6 3.4 95 5 ABP01365 Abp01365 Human ORF 386 6 3.4 131 4 AAB93656 Aab93656 Human pro  
314 6 3.4 96 4 AAU43604 Aau43604 Propionib 387 6 3.4 131 6 ABU33165 Abu33165 Protein e  
315 6 3.4 96 6 ABM40123 Abm40123 Propionib 388 6 3.4 132 4 AAU47392 Aau47392 Propionib  
316 6 3.4 97 4 AAB79810 Aab79810 Corynebac 389 6 3.4 132 5 ABB78773 Abb78773 Corn flav  
317 6 3.4 97 4 AAB79796 Aab79796 Corynebac 390 6 3.4 132 6 ABM64872 Abm64872 Propionib

Abg24258 Novel hum  
Abu06364 Maize Sta  
Add28122 Lymphoma  
Aaw99666 Human sec  
Aao02066 Human pol  
Abb16157 Human ner  
Abp05242 Human ORF  
Abu69154 Human NOV  
Ada45101 Human pol  
Aag04125 Arabidops  
Aag38378 Arabidops  
Aau61535 Propionib  
Abp04030 Human ORF  
Abm58054 Propionib  
Aau68007 Propionib  
Aao17456 Human liv  
Abm64526 Propionib  
Adc31594 Human nov  
Abg60203 Human DIT  
Aag18233 Arabidops  
Abu06365 Maize Sta  
Aab40258 Human ORF  
Aag56021 Arabidops  
Aau66611 Propionib  
Abm63130 Propionib  
Add29464 Mouse cae  
Aag32882 Zea mays  
Aag59688 Zea mays  
Aau65503 Propionib  
Aau31264 Novel hum  
Abm62022 Propionib  
Aab40545 Human ORF  
Aam24343 Human EST  
Abp01461 Human ORF  
Aay73957 Human pro  
Aau40352 Propionib  
Aau30876 Novel hum  
Abm36871 Propionib  
Aag04124 Arabidops  
Aag38377 Arabidops  
Aar25593 Mutant RI  
Abg10851 Novel hum  
Abp76689 Streptomy  
Aam24301 Human EST  
Aau63528 Propionib  
Abg23503 Novel hum  
Abm60047 Propionib  
Aaw74835 Human sec  
Abg95286 Human nov  
Abo34480 Region of  
Aab94940 Human pro  
Aag91805 C glutami  
Aau51169 Propionib  
Abp55524 Human lip  
Abm47688 Propionib  
Abm65163 Propionib  
Aau41835 Propionib  
Abp09027 Human ORF  
Abm38354 Propionib  
Adc94531 E. faeciu  
Aag37466 Arabidops  
Abp64087 Human ORF  
Aag12633 Arabidops  
Aag13266 Arabidops  
Aag34861 Arabidops  
Abb17976 Human ner  
Aau19326 Human G p  
Aam88142 Human imm  
Aab93656 Human pro  
Abu33165 Protein e  
Aau47392 Propionib  
Abb78773 Corn flav  
Abm64872 Propionib

391	6	3.4	132	6	ABM43911	Abm43911	Propionib	464	6	3.4	167	4	ABG01536	Abg01536	Novel hum
392	6	3.4	132	6	ABM64866	Abm64866	Propionib	465	6	3.4	170	4	AAG82430	Aag82430	S. epider
393	6	3.4	132	7	ADB64317	Adb64317	Human pro	466	6	3.4	171	7	ADE38219	Ade38219	Mutant ae
394	6	3.4	133	2	AAR23830	Aar23830	Chicken a	467	6	3.4	171	7	ADE38220	Ade38220	Mutant ae
395	6	3.4	133	4	ABB65017	Abb65017	Drosophil	468	6	3.4	173	6	ABU43050	Abu43050	Protein e
396	6	3.4	133	7	ADC31723	Adc31723	Human nov	469	6	3.4	174	4	AAU23062	Aau23062	Novel hum
397	6	3.4	134	4	AAM80078	Aam80078	Human pro	470	6	3.4	174	4	AAG89144	Aag89144	Human sec
398	6	3.4	134	4	AAM80079	Aam80079	Human pro	471	6	3.4	174	6	ABU11050	Abu11050	Human pro
399	6	3.4	136	3	AAG26066	Aag26066	Zea mays	472	6	3.4	174	6	ABU11051	Abu11051	Human pro
400	6	3.4	136	5	ABP08544	Abp08544	Human ORF	473	6	3.4	175	4	AAU23627	Aau23627	Novel hum
401	6	3.4	137	3	AAB19713	Aab19713	Rat profi	474	6	3.4	175	4	AAO01946	Aao01946	Human pol
402	6	3.4	137	3	AAG19471	Aag19471	Arabidops	475	6	3.4	176	5	ABP65604	Abp65604	Bifidobac
403	6	3.4	137	5	ABP63892	Abp63892	Human ORF	476	6	3.4	176	6	ABU54897	Abu54897	Metabolic
404	6	3.4	138	6	ABU36007	Abu36007	Protein e	477	6	3.4	178	2	AAU39044	Aay39044	M. tuberc
405	6	3.4	139	7	ADC08002	Adc08002	Rice prot	478	6	3.4	178	2	AAU39187	Aay39187	M. tuberc
406	6	3.4	140	2	AAR66395	Aar66395	Human psy	479	6	3.4	179	4	AAG81321	Aag81321	Human AFP
407	6	3.4	140	4	AAO05812	Aao05812	Human pol	480	6	3.4	180	4	ABG14843	Abg14843	Novel hum
408	6	3.4	140	4	AAU67683	Aau67683	Propionib	481	6	3.4	180	4	ABG12827	Abg12827	Novel hum
409	6	3.4	140	4	AAG70862	Aag70862	C albican	482	6	3.4	182	2	AAY60131	Aay60131	Human end
410	6	3.4	140	5	ABB53803	Abb53803	Lactococc	483	6	3.4	182	3	AAB40845	Aab40845	Human ORF
411	6	3.4	140	6	ABM64202	Abm64202	Propionib	484	6	3.4	182	3	AAU56905	Aau56905	Propionib
412	6	3.4	140	7	ADB64647	Adb64647	Human pro	485	6	3.4	182	6	ABM53424	Abm53424	Propionib
413	6	3.4	141	4	AAU56701	Aau56701	Propionib	486	6	3.4	182	6	ADA36181	Ada36181	Acinetoba
414	6	3.4	141	6	ABM53220	Abm53220	Propionib	487	6	3.4	185	4	AAU25668	Aau25668	G protein
415	6	3.4	142	4	AAM25261	Aam25261	Human pro	488	6	3.4	186	3	AAG30372	Aag30372	Arabidops
416	6	3.4	143	5	AAE17483	Aae17483	Human leu	489	6	3.4	186	4	ABB61200	Abb61200	Drosophil
417	6	3.4	143	5	AAU83169	Aau83169	Novel sec	490	6	3.4	187	3	AAU75561	Aay75561	Neisseria
418	6	3.4	144	2	AAU12898	Aay12898	Human 5'	491	6	3.4	187	3	AAG23877	Aag23877	Arabidops
419	6	3.4	144	3	AAU52151	Aay52151	M. paratu	492	6	3.4	188	6	ABU20599	Abu20599	Protein e
420	6	3.4	144	4	AAG67759	Aag67759	Amino aci	493	6	3.4	189	2	AAR65911	Aar65911	Recombina
421	6	3.4	144	4	AAU40042	Aau40042	Propionib	494	6	3.4	189	2	AAW05059	Aaw05059	Aequorin
422	6	3.4	144	4	AAU66505	Aau66505	Propionib	495	6	3.4	189	2	AAW05058	Aaw05058	Aequorin
423	6	3.4	144	6	ABM63024	Abm63024	Propionib	496	6	3.4	189	2	AAW05057	Aaw05057	Aequorin
424	6	3.4	144	6	ABM36561	Abm36561	Propionib	497	6	3.4	189	2	AAW79061	Aaw79061	Apoaequor
425	6	3.4	145	3	AAG13265	Aag13265	Arabidops	498	6	3.4	189	2	AAW62539	Aaw62539	Mutant ap
426	6	3.4	146	2	AAU59845	Aau59845	Human nor	499	6	3.4	189	2	AAW62538	Aaw62538	Wild type
427	6	3.4	146	2	AAW73430	Aaw73430	Human sec	500	6	3.4	189	2	AAW44736	Aaw44736	Apo-aequo
428	6	3.4	146	2	AAW73413	Aaw73413	Human sec	501	6	3.4	189	4	AAB51171	Aab51171	Recombina
429	6	3.4	146	4	ABG01959	Abg01959	Novel hum	502	6	3.4	189	4	AAB51167	Aab51167	Aequorin
430	6	3.4	147	3	AAG12632	Aag12632	Arabidops	503	6	3.4	189	4	ABG20773	Abg20773	Novel hum
431	6	3.4	147	3	AAG34860	Aag34860	Arabidops	504	6	3.4	189	5	ABB82063	Abb82063	Aequorin
432	6	3.4	147	4	AAU53449	Aau53449	Propionib	505	6	3.4	189	5	ABB82067	Abb82067	Recombina
433	6	3.4	147	6	ABM49968	Abm49968	Propionib	506	6	3.4	189	5	AAE13374	Aae13374	Aequorea
434	6	3.4	148	3	AAU97030	Aay97030	Caspase 8	507	6	3.4	189	5	AAE13378	Aae13378	Aequorea
435	6	3.4	150	6	ABU34084	Abu34084	Protein e	508	6	3.4	189	5	ABB82049	Abb82049	Aequorin
436	6	3.4	150	7	ADC88373	Adc88373	Ribosomal	509	6	3.4	189	5	ABB82053	Abb82053	Recombina
437	6	3.4	153	5	ABP33830	Abp33830	Human tra	510	6	3.4	189	5	AAM47869	Aam47869	Aequorea
438	6	3.4	154	3	AAAB42205	Aab42205	Human ORF	511	6	3.4	189	5	ABB05763	Abb05763	Human G p
439	6	3.4	154	3	AAG32663	Aag32663	Zea mays	512	6	3.4	189	7	ADE39697	Ade39697	Wild type
440	6	3.4	157	3	AAG30374	Aag30374	Arabidops	513	6	3.4	190	4	ABB64581	Abb64581	Drosophil
441	6	3.4	157	4	AAG72851	Aag72851	Human olf	514	6	3.4	191	3	AAG16807	Aag16807	Arabidops
442	6	3.4	157	4	AAO04329	Aao04329	Human pol	515	6	3.4	191	3	AAG09704	Aag09704	Arabidops
443	6	3.4	157	6	ABU31970	Abu31970	Protein e	516	6	3.4	191	3	AAG45587	Aag45587	Arabidops
444	6	3.4	158	4	ABG28649	Abg28649	Novel hum	517	6	3.4	191	3	AAG45584	Aag45584	Arabidops
445	6	3.4	158	6	ABM64847	Abm64847	Propionib	518	6	3.4	191	4	AAM95494	Aam95494	Human rep
446	6	3.4	159	4	AAU22264	Aau22264	Human car	519	6	3.4	191	4	AAG70792	Aag70792	S cerevis
447	6	3.4	159	7	ADE46232	Ade46232	Human car	520	6	3.4	191	4	AAG70729	Aag70729	S cerevis
448	6	3.4	160	3	ABM53652	Aab53652	Human col	521	6	3.4	191	4	ABB96178	Abb96178	Human tes
449	6	3.4	160	4	AAM40845	Aam40845	Human pol	522	6	3.4	192	6	ABU17334	Abu17334	Protein e
450	6	3.4	160	4	AAU50012	Aau50012	Propionib	523	6	3.4	192	7	ADE39698	Ade39698	Cysteine
451	6	3.4	160	4	AAU43406	Aau43406	Propionib	524	6	3.4	194	6	ABU19797	Abu19797	Protein e
452	6	3.4	160	6	ABM46531	Abm46531	Propionib	525	6	3.4	195	4	ABG21534	Abg21534	Novel hum
453	6	3.4	160	6	ABM39925	Abm39925	Propionib	526	6	3.4	195	5	ABB06182	Abb06182	Green flu
454	6	3.4	161	3	AAG37465	Aag37465	Arabidops	527	6	3.4	196	1	AAP60795	Aap60795	Aquarin p
455	6	3.4	161	4	AAU67944	Aau67944	Propionib	528	6	3.4	196	1	AAP60273	Aap60273	Sequence
456	6	3.4	161	6	ABM64463	Abm64463	Propionib	529	6	3.4	196	2	AAR04822	Aar04822	Aequorin
457	6	3.4	162	3	AAG19470	Aag19470	Arabidops	530	6	3.4	196	2	AAR65914	Aar65914	Recombina
458	6	3.4	163	3	AAG30373	Aag30373	Arabidops	531	6	3.4	196	2	AAR65910	Aar65910	Recombina
459	6	3.4	163	5	ABB04854	Abb04854	LDL recep	532	6	3.4	196	2	AAR65913	Aar65913	Recombina
460	6	3.4	164	4	ABG21128	Abg21128	Novel hum	533	6	3.4	196	2	AAR65912	Aar65912	Recombina
461	6	3.4	166	5	AAE25080	Aae25080	Mouse ost	534	6	3.4	196	2	AAR77875	Aar77875	Apoaequor
462	6	3.4	167	2	AAW28017	Aaw28017	Staphyloc	535	6	3.4	196	2	AAW05055	Aaw05055	Apoaequor
463	6	3.4	167	3	AAG23878	Aag23878	Arabidops	536	6	3.4	196	2	AAW05054	Aaw05054	Apoaequor



537	6	3.4	196	2	AAW05056	Aaw05056	Apoaequor	610	6	3.4	199	6	ADB25545	Adb25545	Mutant je
538	6	3.4	196	2	AAW57996	Aaw57996	Apoaequor	611	6	3.4	199	6	ADB25583	Adb25583	Mutant je
539	6	3.4	196	2	AAW57997	Aaw57997	Apoaequor	612	6	3.4	199	6	ADB25481	Adb25481	Mutant je
540	6	3.4	196	2	AAW59921	Aaw59921	Apoaequor	613	6	3.4	199	6	ADB25543	Adb25543	Mutant je
541	6	3.4	196	2	AAW62540	Aaw62540	Wild type	614	6	3.4	199	6	ADB25580	Adb25580	Mutant je
542	6	3.4	196	2	AAW44737	Aaw44737	Wild type	615	6	3.4	199	6	ADB25579	Adb25579	Mutant je
543	6	3.4	196	4	AAB51168	Aab51168	Aequorin	616	6	3.4	199	6	ADB25475	Adb25475	Mutant je
544	6	3.4	196	4	AAB51165	Aab51165	Recombina	617	6	3.4	199	6	ADB25547	Adb25547	Mutant je
545	6	3.4	196	4	AAB51166	Aab51166	Recombina	618	6	3.4	199	6	ADB25537	Adb25537	Mutant je
546	6	3.4	196	4	AAB51169	Aab51169	Aequorin	619	6	3.4	199	6	ABM58119	Abm58119	Propionib
547	6	3.4	196	4	AAB51163	Aab51163	Apoaequor	620	6	3.4	200	3	AAG53753	Aag53753	Arabidops
548	6	3.4	196	4	AAB51170	Aab51170	Aequorin	621	6	3.4	200	4	AAU47538	Aau47538	Propionib
549	6	3.4	196	4	AAB51164	Aab51164	Recombina	622	6	3.4	200	6	ABM44057	Abm44057	Propionib
550	6	3.4	196	5	ABB82059	Abb82059	Apoaequor	623	6	3.4	201	4	AAU59473	Aau59473	Propionib
551	6	3.4	196	5	ABB82064	Abb82064	Aequorin	624	6	3.4	201	6	ABM55992	Abm55992	Propionib
552	6	3.4	196	5	ABB82066	Abb82066	Aequorin	625	6	3.4	203	5	AAU09673	Aau09673	Mycobacte
553	6	3.4	196	5	ABB82060	Abb82060	Recombina	626	6	3.4	203	6	ABU19459	Abu19459	Protein e
554	6	3.4	196	5	ABB82062	Abb82062	Recombina	627	6	3.4	203	6	ABU11506	Abu11506	Human MDD
555	6	3.4	196	5	ABB82061	Abb82061	Recombina	628	6	3.4	203	6	ADA35784	Ada35784	Acinetoba
556	6	3.4	196	5	ABB82065	Abb82065	Aequorin	629	6	3.4	205	2	AAR41022	Aar41022	Protein h
557	6	3.4	196	5	AAE13372	Aae13372	Aequorea	630	6	3.4	205	5	ABB78997	Abb78997	Human HSB
558	6	3.4	196	5	AAE13377	Aae13377	Aequorea	631	6	3.4	205	5	AAO18753	Aao18753	Human HSP
559	6	3.4	196	5	AAE13373	Aae13373	Aequorea	632	6	3.4	205	7	ADE83437	Ade83437	Human Pro
560	6	3.4	196	5	AAE13376	Aae13376	Aequorea	633	6	3.4	205	7	ADE75363	Ade75363	Human SHS
561	6	3.4	196	5	AAE13370	Aae13370	Aequorea	634	6	3.4	206	2	AAy60434	Aay60434	Human nor
562	6	3.4	196	5	AAE13371	Aae13371	Aequorea	635	6	3.4	207	4	ABB68796	Abb68796	Drosophil
563	6	3.4	196	5	AAE13371	Aae13371	Aequorea	636	6	3.4	208	4	AB925559	Aab92559	Human pro
564	6	3.4	196	5	ABB82048	Abb82048	Recombina	637	6	3.4	208	4	ABG04806	Abg04806	Novel hum
565	6	3.4	196	5	ABB82050	Abb82050	Aequorin	638	6	3.4	208	6	ABU19621	Abu19621	Protein e
566	6	3.4	196	5	ABB82051	Abb82051	Aequorin	639	6	3.4	209	2	AAy50007	Aay50007	Thermus t
567	6	3.4	196	5	ABB82052	Abb82052	Aequorin	640	6	3.4	209	4	AAW08000	Aam08000	Human bon
568	6	3.4	196	5	ABB82045	Abb82045	Apoaequor	641	6	3.4	210	2	AAW85052	Aaw85052	Epsilon s
569	6	3.4	196	5	ABB82047	Abb82047	Recombina	642	6	3.4	210	2	AAW44738	Aaw44738	Chimeric
570	6	3.4	196	6	ABP58058	Abp58058	Aequorea	643	6	3.4	210	2	AAW44739	Aaw44739	Mutant ap
571	6	3.4	196	7	ADE38221	Ade38221	Mutant ae	644	6	3.4	210	4	ABG12907	Abgl12907	Novel hum
572	6	3.4	197	4	ABG17061	Abg17061	Novel hum	645	6	3.4	215	3	AAG09703	Aag09703	Arabidops
573	6	3.4	197	5	AAE13375	Aae13375	Aequorea	646	6	3.4	215	3	AAG26454	Aag26454	Arabidops
574	6	3.4	198	4	AAE13375	Aae13375	Aequorea	647	6	3.4	215	3	AAG16806	Aag16806	Arabidops
575	6	3.4	198	6	ADB25565	Adb25565	Mutant je	648	6	3.4	215	3	AAG45590	Aag45590	Arabidops
576	6	3.4	198	6	ADB25507	Adb25507	Mutant je	649	6	3.4	215	3	AAG43524	Aag43524	Arabidops
577	6	3.4	198	6	ADB25595	Adb25595	Mutant je	650	6	3.4	215	3	AAG45583	Aag45583	Arabidops
578	6	3.4	198	6	ADB25594	Adb25594	Mutant je	651	6	3.4	215	3	AAG28817	Aag28817	Arabidops
579	6	3.4	198	6	ADB25596	Adb25596	Mutant je	652	6	3.4	215	3	AAG45586	Aag45586	Arabidops
580	6	3.4	198	6	ADB25505	Adb25505	Wild-type	653	6	3.4	215	4	ABB69711	Abb69711	Drosophil
581	6	3.4	198	6	ADB25567	Adb25567	Mutant je	654	6	3.4	216	4	AAg72817	Aag72817	Human olf
582	6	3.4	198	6	ADB25515	Adb25515	Mutant je	655	6	3.4	216	4	AAW80056	Aam80056	Human pro
583	6	3.4	198	6	ADB25598	Adb25598	Mutant je	656	6	3.4	216	4	AAW35238	Aay35238	Amino aci
584	6	3.4	198	6	ADB25511	Adb25511	Mutant je	657	6	3.4	217	2	AAy35238	Aay35238	DNA clone
585	6	3.4	198	6	ADB25592	Adb25592	Mutant je	658	6	3.4	217	7	ABW73881	Abw73881	MyokL pro
586	6	3.4	198	6	ADB25519	Adb25519	Mutant je	659	6	3.4	218	5	AAE15442	Aae15442	Human dru
587	6	3.4	198	6	ADB25517	Adb25517	Mutant je	660	6	3.4	219	3	AAG43424	Aag43424	Arabidops
588	6	3.4	198	6	ADB25597	Adb25597	Mutant je	661	6	3.4	219	3	AAG06202	Aag06202	Arabidops
589	6	3.4	198	6	ADB25513	Adb25513	Mutant je	662	6	3.4	220	5	ABP25878	Abp25878	Streptoco
590	6	3.4	198	6	ADB25561	Adb25561	Mutant je	663	6	3.4	222	4	AAU68588	Aau68588	Human nov
591	6	3.4	198	6	ADB25563	Adb25563	Mutant je	664	6	3.4	222	6	ADA83789	Ada83789	Human HOX
592	6	3.4	198	6	ADB25509	Adb25509	Mutant je	665	6	3.4	226	7	ADC07882	Adc07882	Rice prot
593	6	3.4	198	6	ADB25559	Adb25559	Mutant je	666	6	3.4	227	6	ABU96685	Abu96685	Human nuc
594	6	3.4	198	6	ADB25593	Adb25593	Mutant je	667	6	3.4	227	6	ABU62625	Abu62625	Pertussis
595	6	3.4	198	6	ADB94787	Adb94787	Programme	668	6	3.4	228	3	AAg19469	Aag19469	Arabidops
596	6	3.4	199	3	AAG42936	Aag42936	Arabidops	669	6	3.4	228	5	ABU65189	Abu65189	Human NOV
597	6	3.4	199	4	AAU61600	Aau61600	Propionib	670	6	3.4	229	3	AAg16687	Aag16687	Arabidops
598	6	3.4	199	6	ADB25487	Adb25487	Mutant je	671	6	3.4	229	6	ABU28068	Abu28068	Protein e
599	6	3.4	199	6	ADB25485	Adb25485	Mutant je	672	6	3.4	230	3	AAG54552	Aag54552	Zea may
600	6	3.4	199	6	ADB25582	Adb25582	Mutant je	673	6	3.4	230	4	AAW93596	Aam93596	Human pol
601	6	3.4	199	6	ADB25584	Adb25584	Mutant je	674	6	3.4	230	5	ABP30805	Abp30805	Streptoco
602	6	3.4	199	6	ADB25483	Adb25483	Mutant je	675	6	3.4	230	5	ABP26796	Abp26796	Streptoco
603	6	3.4	199	6	ADB25541	Adb25541	Mutant je	676	6	3.4	230	7	ADC56779	Adc56779	Splice va
604	6	3.4	199	6	ADB25473	Adb25473	Wild-type	677	6	3.4	231	3	AAg16686	Aag16686	Arabidops
605	6	3.4	199	6	ADB25477	Adb25477	Mutant je	678	6	3.4	231	3	AAG23599	Aag23599	Arabidops
606	6	3.4	199	6	ADB25479	Adb25479	Mutant je	679	6	3.4	233	3	AAG24668	Aag24668	Arabidops
607	6	3.4	199	6	ADB25578	Adb25578	Mutant je	680	6	3.4	233	3	AAG35484	Aag35484	Arabidops
608	6	3.4	199	6	ADB25539	Adb25539	Mutant je	681	6	3.4	233	3	AAG50339	Aag50339	Arabidops
609	6	3.4	199	6	ADB25581	Adb25581	Mutant je	682	6	3.4	234	3	AAG51358	Aag51358	Arabidops

683	6	3.4	234	3	AAG18047	Aag18047 Arabidops	756	6	3.4	271	4	AAM18360	Aam18360 Peptide #
684	6	3.4	234	6	ABU29658	Abu29658 Protein e	757	6	3.4	271	4	ABB65959	Abb65959 Drosophil
685	6	3.4	234	7	ABM74452	Abm74452 DNA clone	758	6	3.4	271	4	ABB37857	Abb37857 Peptide #
686	6	3.4	235	4	AAU67819	Aau67819 Propionib	759	6	3.4	271	4	AAM31263	Aam31263 Peptide #
687	6	3.4	235	5	ABU05899	Abu05899 M. tuberc	760	6	3.4	271	4	ABB32143	Abb32143 Peptide #
688	6	3.4	235	6	ABM64338	Abm64338 Propionib	761	6	3.4	271	4	ABB22688	Abb22688 Protein #
689	6	3.4	235	7	ADB65309	Adb65309 Human pro	762	6	3.4	271	4	AAM70986	Aam70986 Human bon
690	6	3.4	237	4	AAM79094	Aam79094 Human pro	763	6	3.4	271	4	AAM58487	Aam58487 Human bra
691	6	3.4	237	7	ADB64930	Adb64930 Human pro	764	6	3.4	271	4	ABG52705	Abg52705 Human liv
692	6	3.4	238	3	AAG53752	Aag53752 Arabidops	765	6	3.4	271	5	ABG40777	Abg40777 Human pep
693	6	3.4	238	4	AAU17469	Aau17469 Novel sig	766	6	3.4	272	6	ABU38840	Abu38840 Protein e
694	6	3.4	238	4	ABB67392	Abb67392 Drosophil	767	6	3.4	273	4	ABG71376	Aag71376 Human sec
695	6	3.4	238	7	ADB94177	Adb94177 Human nov	768	6	3.4	275	4	AAB78965	Aab78965 C. glutam
696	6	3.4	238	7	ADC95201	Adc95201 E. faeciu	769	6	3.4	275	4	AAG90648	Aag90648 C. glutami
697	6	3.4	239	6	ABU33874	Abu33874 Protein e	770	6	3.4	275	5	ABB49829	Abb49829 Listeria
698	6	3.4	242	2	AAU34764	Aay34764 Chlamydia	771	6	3.4	275	6	ABU21540	Abu21540 Protein e
699	6	3.4	242	6	ABM65634	Abm65634 Propionib	772	6	3.4	275	6	ADA89397	Ada89397 Wheat hyp
700	6	3.4	243	3	AAG06092	Aag06092 Arabidops	773	6	3.4	276	3	AAG43423	Aag43423 Arabidops
701	6	3.4	244	3	AAG45153	Aag45153 Arabidops	774	6	3.4	276	3	AAG06201	Aag06201 Arabidops
702	6	3.4	244	5	AAE25079	Aae25079 Mouse ost	775	6	3.4	277	6	ADA89395	Ada89395 Wheat hyp
703	6	3.4	246	3	AAG45152	Aag45152 Arabidops	776	6	3.4	278	3	AAU53923	Aay53923 A homolog
704	6	3.4	246	4	AAG92104	Aag92104 C glutami	777	6	3.4	278	5	ABB90722	Abb90722 Human Tum
705	6	3.4	246	4	ABG29264	Abg29264 Novel hum	778	6	3.4	278	5	ABU05355	Abu05355 M. tuberc
706	6	3.4	246	4	ABG11394	Abg11394 Novel hum	779	6	3.4	278	6	ABU54429	Abu54429 Human tum
707	6	3.4	247	4	AAE11934	Aae11934 Human CG2	780	6	3.4	279	7	ADC56780	Adc56780 Splice va
708	6	3.4	247	7	ADB64636	Adb64636 Human pro	781	6	3.4	280	2	AAU42694	Aay42694 Rat kd312
709	6	3.4	249	6	ADA09945	Ada09945 Rat hypot	782	6	3.4	280	2	AAU05317	Aay05317 Human sec
710	6	3.4	249	7	ADC46749	Adc46749 Thalecres	783	6	3.4	280	3	AAU66728	Aay66728 Membrane-
711	6	3.4	249	7	ADD30406	Add30406 Plant yie	784	6	3.4	280	3	AAB42619	Aab42619 Human ORF
712	6	3.4	249	7	ADE31559	Ade31559 Plant yie	785	6	3.4	280	4	AAM93544	Aam93544 Human pol
713	6	3.4	250	3	AAG45582	Aag45582 Arabidops	786	6	3.4	280	4	AAB50959	Aab50959 Human PRO
714	6	3.4	250	3	AAG16805	Aag16805 Arabidops	787	6	3.4	280	4	AAU12400	Aau12400 Human PRO
715	6	3.4	250	3	AAG26453	Aag26453 Arabidops	788	6	3.4	280	4	AAU29324	Aau29324 Human PRO
716	6	3.4	250	3	AAG43523	Aag43523 Arabidops	789	6	3.4	280	4	AAU62348	Aau62348 Propionib
717	6	3.4	250	3	AAG45585	Aag45585 Arabidops	790	6	3.4	280	4	AAB65251	Aab65251 Human PRO
718	6	3.4	250	3	AAG45589	Aag45589 Arabidops	791	6	3.4	280	4	AAE11938	Aae11938 Human lip
719	6	3.4	250	3	AAG09702	Aag09702 Arabidops	792	6	3.4	280	4	AAE11932	Aae11932 Human CG2
720	6	3.4	250	5	ABB90903	Abb90903 Herbicida	793	6	3.4	280	5	ABB84899	Abb84899 Human PRO
721	6	3.4	251	3	AAG16136	Aag16136 Arabidops	794	6	3.4	280	5	ABP64809	Abp64809 Human PRO
722	6	3.4	251	3	AAG42935	Aag42935 Arabidops	795	6	3.4	280	5	ABP64809	Abp64809 Human PRO
723	6	3.4	253	3	AAG23598	Aag23598 Arabidops	796	6	3.4	280	5	ABB95505	Abb95505 Human ang
724	6	3.4	254	3	AAG43522	Aag43522 Arabidops	797	6	3.4	280	6	ABU58700	Abu58700 Human PRO
725	6	3.4	254	5	ABG66666	Abg66666 Human nov	798	6	3.4	280	6	ABU88248	Abu88248 Novel hum
726	6	3.4	255	7	ADB65722	Adb65722 Human pro	799	6	3.4	280	6	ABU84563	Abu84563 Human sec
727	6	3.4	256	3	AAG18046	Aag18046 Arabidops	800	6	3.4	280	6	ABR66437	Abr66437 Human sec
728	6	3.4	256	3	AAG51357	Aag51357 Arabidops	801	6	3.4	280	6	ABR65827	Abr65827 Human sec
729	6	3.4	256	4	ABG19812	Abg19812 Novel hum	802	6	3.4	280	6	ABU99767	Abu99767 Human PRO
730	6	3.4	257	3	AAU91302	Aay91302 Group B S	803	6	3.4	280	6	ABU58066	Abu58066 Human PRO
731	6	3.4	257	6	ABU38808	Abu38808 Protein e	804	6	3.4	280	6	ABU59144	Abu59144 Novel hum
732	6	3.4	260	4	AAG90297	Aag90297 C glutami	805	6	3.4	280	6	ABU82656	Abu82656 Human sec
733	6	3.4	261	3	AAB43456	Aab43456 Human can	806	6	3.4	280	6	ABU83006	Abu83006 Human PRO
734	6	3.4	261	6	ABM67359	Abm67359 Photorhab	807	6	3.4	280	6	ABO17844	Abol17844 Novel hum
735	6	3.4	263	6	ABU22022	Abu22022 Protein e	808	6	3.4	280	6	ABR68376	Abr68376 Human sec
736	6	3.4	265	4	AAU59807	Aau59807 Propionib	809	6	3.4	280	6	ABU60575	Abu60575 Human sec
737	6	3.4	265	6	ABM56326	Abm56326 Propionib	810	6	3.4	280	6	ABU96429	Abu96429 Novel hum
738	6	3.4	265	6	ABU96727	Abu96727 Human nuc	811	6	3.4	280	6	ABU92860	Abu92860 Human sec
739	6	3.4	266	5	ABP30450	Abp30450 Streptoco	812	6	3.4	280	6	ABO08937	Abol08937 Human sec
740	6	3.4	266	5	ABB90733	Abb90733 Human Tum	813	6	3.4	280	6	ABO02989	Abol02989 Human sec
741	6	3.4	266	5	ABB90781	Abb90781 Mouse Tum	814	6	3.4	280	6	ABR75143	Abr75143 Human sec
742	6	3.4	266	5	ABG70951	Abg70951 Human Ras	815	6	3.4	280	6	ABR94905	Abr94905 Human sec
743	6	3.4	266	6	ABU54488	Abu54488 Mouse tum	816	6	3.4	280	6	ABU13957	Abul13957 Human PRO
744	6	3.4	266	6	ABU54440	Abu54440 Human tum	817	6	3.4	280	6	ABU85878	Abu85878 Human PRO
745	6	3.4	266	6	ABU25500	Abu25500 Protein e	818	6	3.4	280	6	ABU99038	Abu99038 Novel hum
746	6	3.4	267	3	AAG27467	Aag27467 Arabidops	819	6	3.4	280	6	ABU98253	Abu98253 Novel hum
747	6	3.4	267	3	AAG17403	Aag17403 Arabidops	820	6	3.4	280	6	ABU81098	Abu81098 Human PRO
748	6	3.4	268	4	AAG91104	Aag91104 C glutami	821	6	3.4	280	6	ABU91959	Abu91959 Novel hum
749	6	3.4	268	6	ABO53572	Abol53572 Novel hum	822	6	3.4	280	6	ABU89652	Abu89652 Human PRO
750	6	3.4	269	4	AAM40145	Aam40145 Human pol	823	6	3.4	280	6	ABU86493	Abu86493 Human sec
751	6	3.4	269	5	ABB92448	Abb92448 Herbicida	824	6	3.4	280	6	ABU67706	Abu67706 Human sec
752	6	3.4	269	5	AAU80184	Aau80184 Embryonic	825	6	3.4	280	6	ABU80734	Abu80734 Human PRO
753	6	3.4	270	3	AAU93489	Aay93489 Amino aci	826	6	3.4	280	6	ABU72542	Abu72542 Novel hum
754	6	3.4	270	5	ABG60609	Abg60609 Rat potas	827	6	3.4	280	6	ABR99652	Abr99652 Human sec
755	6	3.4	270	6	ADB10356	Adb10356 Alloiococ	828	6	3.4	280	6	ABR99042	Abr99042 Human sec



829	6	3.4	280	6	ABO16565	Abol16565 Human sec	902	6	3.4	280	6	ABU94892	Abu94892 Human PRO
830	6	3.4	280	6	ABR92465	Abr92465 Human sec	903	6	3.4	280	6	ABO04819	Abo04819 Human PRO
831	6	3.4	280	6	ABO19106	Abol19106 Human sec	904	6	3.4	280	6	ABR70568	Abr70568 Human sec
832	6	3.4	280	6	ABR78527	Abr78527 Human sec	905	6	3.4	280	6	ABU92375	Abu92375 Novel hum
833	6	3.4	280	6	ABU71425	Abu71425 Human neo	906	6	3.4	280	6	ABU98733	Abu98733 Human PRO
834	6	3.4	280	6	ABU85263	Abu85263 Novel hum	907	6	3.4	280	6	ABR66132	Abr66132 Human sec
835	6	3.4	280	6	ABU66798	Abu66798 Human PRO	908	6	3.4	280	6	ABR64849	Abr64849 Human sec
836	6	3.4	280	6	ABO00402	Abol00402 Novel hum	909	6	3.4	280	6	ABU59440	Abu59440 Novel hum
837	6	3.4	280	6	ABM58867	Abm58867 Propionib	910	6	3.4	280	6	ABU79774	Abu79774 Human PRO
838	6	3.4	280	6	ABO11734	Abol11734 Human sec	911	6	3.4	280	6	ABU67074	Abu67074 Human sec
839	6	3.4	280	6	ABO02379	Abol02379 Human sec	912	6	3.4	280	6	ABU93165	Abu93165 Human sec
840	6	3.4	280	6	ABU88953	Abu88953 Novel hum	913	6	3.4	280	6	ABU96124	Abu96124 Human PRO
841	6	3.4	280	6	ABU83648	Abu83648 Human sec	914	6	3.4	280	6	ABU91344	Abu91344 Novel hum
842	6	3.4	280	6	ABO06449	Abol06449 Novel hum	915	6	3.4	280	6	ABU90437	Abu90437 Novel hum
843	6	3.4	280	6	ABR59485	Abr59485 Human sec	916	6	3.4	280	6	ABO09852	Abo09852 Human sec
844	6	3.4	280	6	ABO09547	Abol09547 Human sec	917	6	3.4	280	6	ABO11124	Abo11124 Human sec
845	6	3.4	280	6	ABO19411	Abol19411 Novel hum	918	6	3.4	280	6	ABR71178	Abr71178 Human sec
846	6	3.4	280	6	ABO11429	Abol11429 Human sec	919	6	3.4	280	6	ABU87786	Abu87786 Human PRO
847	6	3.4	280	6	ABR67047	Abr67047 Human sec	920	6	3.4	280	6	ABU91654	Abu91654 Human PRO
848	6	3.4	280	6	ABO16260	Abol16260 Human sec	921	6	3.4	280	6	ABU84868	Abu84868 Human sec
849	6	3.4	280	6	ABO13966	Abol13966 Human sec	922	6	3.4	280	6	ABR69958	Abr69958 Human sec
850	6	3.4	280	6	ABU65869	Abu65869 Human sec	923	6	3.4	280	6	ABU80335	Abu80335 Human PRO
851	6	3.4	280	6	ABO07717	Abol07717 Human PRO	924	6	3.4	280	6	ABU92206	Abu92206 Novel hum
852	6	3.4	280	6	ABO03904	Abol03904 Human sec	925	6	3.4	280	6	ABU93604	Abu93604 Human PRO
853	6	3.4	280	6	ABR67352	Abr67352 Human sec	926	6	3.4	280	6	ABO10157	Abo10157 Human sec
854	6	3.4	280	6	ABO15955	Abol15955 Human sec	927	6	3.4	280	6	ABO09242	Abo09242 Human sec
855	6	3.4	280	6	ABU59879	Abu59879 Novel sec	928	6	3.4	280	6	ABU10912	Abu10912 Human PRO
856	6	3.4	280	6	ABU56236	Abu56236 Human sec	929	6	3.4	280	6	ABU10810	Abu10810 Human sec
857	6	3.4	280	6	ABU65564	Abu65564 Human PRO	930	6	3.4	280	6	ABU81664	Abu81664 Novel hum
858	6	3.4	280	6	ABU95509	Abu95509 Novel hum	931	6	3.4	280	6	ABU95819	Abu95819 Human PRO
859	6	3.4	280	6	ABU71412	Abu71412 Human PRO	932	6	3.4	280	6	ABU97028	Abu97028 Novel hum
860	6	3.4	280	6	ABO08022	Abol08022 Human PRO	933	6	3.4	280	6	ABR70873	Abr70873 Human sec
861	6	3.4	280	6	ABR70263	Abr70263 Human sec	934	6	3.4	280	6	ABO05224	Abo05224 Novel hum
862	6	3.4	280	6	ABR69596	Abr69596 Human sec	935	6	3.4	280	6	ABO08632	Abo08632 Human sec
863	6	3.4	280	6	ABO01737	Abol01737 Human PRO	936	6	3.4	280	6	ABU88603	Abu88603 Human sec
864	6	3.4	280	6	ABU81539	Abu81539 Human PRO	937	6	3.4	280	6	ABO34117	Abo34117 Human PRO
865	6	3.4	280	6	ABR60336	Abr60336 Human sec	938	6	3.4	280	6	ABO05839	Abo05839 Human sec
866	6	3.4	280	6	ABR68071	Abr68071 Human sec	939	6	3.4	280	6	ABR74228	Abr74228 Human sec
867	6	3.4	280	6	ABR65459	Abr65459 Human sec	940	6	3.4	280	6	ABR95820	Abr95820 Human sec
868	6	3.4	280	6	ABR68681	Abr68681 Human sec	941	6	3.4	280	6	ABR81117	Abr81117 Human sec
869	6	3.4	280	6	ABR72093	Abr72093 Human sec	942	6	3.4	280	6	ABR81422	Abr81422 Human sec
870	6	3.4	280	6	ABU59291	Abu59291 Human sec	943	6	3.4	280	6	ABM01118	Abm01118 Human sec
871	6	3.4	280	6	ABU85573	Abu85573 Human PRO	944	6	3.4	280	6	ABR88720	Abr88720 Human sec
872	6	3.4	280	6	ABU89263	Abu89263 Human sec	945	6	3.4	280	6	ADA45977	Ada45977 Novel hum
873	6	3.4	280	6	ABU83343	Abu83343 Human sec	946	6	3.4	280	6	ABM77541	Abm77541 Human sec
874	6	3.4	280	6	ABU95199	Abu95199 Novel hum	947	6	3.4	280	6	Abo29025	Abo29025 Human sec
875	6	3.4	280	6	ABU90747	Abu90747 Novel hum	948	6	3.4	280	6	ABO31770	Abo31770 Human sec
876	6	3.4	280	6	ABU84258	Abu84258 Human sec	949	6	3.4	280	6	ABM08187	Abm08187 Human sec
877	6	3.4	280	6	ABU93909	Abu93909 Novel hum	950	6	3.4	280	6	ABO40667	Abo40667 Human sec
878	6	3.4	280	6	ABO25988	Abo25988 Human PRO	951	6	3.4	280	6	ABO36092	Abo36092 Human PRO
879	6	3.4	280	6	ABR65154	Abr65154 Human sec	952	6	3.4	280	6	ABO44231	Abo44231 Human PRO
880	6	3.4	280	6	ABR68986	Abr68986 Human sec	953	6	3.4	280	6	ADA78354	Ada78354 Human sec
881	6	3.4	280	6	ABO06802	Abo06802 Human sec	954	6	3.4	280	6	ABM25026	Abm25026 Human sec
882	6	3.4	280	6	ABO25069	Abo25069 Human sec	955	6	3.4	280	6	ADA76408	Ada76408 Human PRO
883	6	3.4	280	6	ABR99347	Abr99347 Human sec	956	6	3.4	280	6	ABO03294	Abo03294 Human sec
884	6	3.4	280	6	ABU57231	Abu57231 Human PRO	957	6	3.4	280	6	ABR90550	Abr90550 Human sec
885	6	3.4	280	6	ABU86183	Abu86183 Novel hum	958	6	3.4	280	6	ABM17464	Abm17464 Human sec
886	6	3.4	280	6	ABU82470	Abu82470 Novel hum	959	6	3.4	280	6	ABR95210	Abr95210 Human sec
887	6	3.4	280	6	ABU87481	Abu87481 Human PRO	960	6	3.4	280	6	ABR95515	Abr95515 Human sec
888	6	3.4	280	6	ABU83953	Abu83953 Human sec	961	6	3.4	280	6	ABO21753	Abo21753 Human sec
889	6	3.4	280	6	ABO08327	Abo08327 Human PRO	962	6	3.4	280	6	ADA19058	Ada19058 Human PRO
890	6	3.4	280	6	ABU82038	Abu82038 Novel hum	963	6	3.4	280	6	ABR98017	Abr98017 Human sec
891	6	3.4	280	6	ABU66202	Abu66202 Novel hum	964	6	3.4	280	6	ABR87805	Abr87805 Human sec
892	6	3.4	280	6	ABR60031	Abr60031 Human sec	965	6	3.4	280	6	ABM77846	Abm77846 Human sec
893	6	3.4	280	6	ABU94219	Abu94219 Novel hum	966	6	3.4	280	6	ABM28076	Abm28076 Human sec
894	6	3.4	280	6	ABO00092	Abo00092 Novel hum	967	6	3.4	280	6	ABM06357	Abm06357 Human sec
895	6	3.4	280	6	ABR66742	Abr66742 Human sec	968	6	3.4	280	6	ABM03863	Abm03863 Human sec
896	6	3.4	280	6	ABR91160	Abr91160 Human sec	969	6	3.4	280	6	ABM35314	Abm35314 Human sec
897	6	3.4	280	6	ABU58997	Abu58997 Human sec	970	6	3.4	280	6	ABM26551	Abm26551 Human sec
898	6	3.4	280	6	ABU94587	Abu94587 Human PRO	971	6	3.4	280	6	ABO48333	Abo48333 Human sec
899	6	3.4	280	6	ABU79469	Abu79469 Human PRO	972	6	3.4	280	6	ADA61681	Ada61681 Homo sapi
900	6	3.4	280	6	ABU86798	Abu86798 Human sec	973	6	3.4	280	6	ADB19466	Adb19466 Novel hum
901	6	3.4	280	6	ABU87103	Abu87103 Novel hum	974	6	3.4	280	6	ADB28007	Adb28007 Human PRO

975	6	3.4	280	6	ADA86486	Novel hum
976	6	3.4	280	6	ADB16050	Human PRO
977	6	3.4	280	6	ABR93075	Human sec
978	6	3.4	280	6	ABO24836	Human sec
979	6	3.4	280	6	ADA37830	Human sec
980	6	3.4	280	6	ABM11847	Human sec
981	6	3.4	280	6	ABM02948	Human sec
982	6	3.4	280	6	ABM16244	Human sec
983	6	3.4	280	6	ABO27805	Human sec
984	6	3.4	280	6	ABM29296	Human sec
985	6	3.4	280	6	ABM07272	Human sec
986	6	3.4	280	6	ABM21366	Human sec
987	6	3.4	280	6	ABM09712	Human sec
988	6	3.4	280	6	ABO41582	Human sec
989	6	3.4	280	6	ABO36397	Human PRO
990	6	3.4	280	6	ABO43926	Human PRO
991	6	3.4	280	6	ADA47836	Human PRO
992	6	3.4	280	6	ABM76626	Human sec
993	6	3.4	280	6	ABM76322	Human sec
994	6	3.4	280	6	ABM25941	Human sec
995	6	3.4	280	6	ABM26246	Human sec
996	6	3.4	280	6	ADA21516	Human sec
997	6	3.4	280	6	ABO03599	Human sec
998	6	3.4	280	6	ABO02684	Human sec
999	6	3.4	280	6	ABR90855	Human sec
1000	6	3.4	280	8	ADD87455	Human PRO

ALIGNMENTS

RESULT 1

AAAY72663  
ID AAAY72663 standard; protein; 174 AA.

AC AAAY72663;

XX 31-MAY-2001 (first entry)

XX Human peptidoglycan recognition protein-related liver protein.

XX Peptidoglycan recognition protein-related liver protein; PGRP-L;  
KW food additive; food preservative; breast cancer; ovarian cancer;  
KW immune disorder; Addison's disease; allergy; cardiovascular disorder;  
KW myocardial ischaemia; wound healing; neurological disease; vasotropic;  
KW Alzheimer's disease; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cerebroprotective; antibacterial;  
KW virucide; fungicide; ophthalmological; human; gene therapy.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FH	Region	12. .18
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	34. .47
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	51. .57
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	64. .75
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	80. .84
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	96. .133
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	135. .137
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"

FT	Region	145. .154
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	161. .167
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"
FT	Region	172. .174
FT		/note= "Antigenic region of human peptidoglycan recognition protein-related liver protein"

XX WO200114545-A1.

PN 01-MAR-2001.

XX 18-AUG-2000; 2000WO-US022877.

XX 20-AUG-1999; 99US-0149715P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Young PE, Rosen CA, Duan RD;

XX WPI; 2001-160115/16.

XX N-PSDB; AAD02742.

XX Isolated nucleic acids encoding human and murine peptidoglycan recognition protein-related liver (PGRP-L) proteins, useful for preventing or treating autoimmune diseases e.g. rheumatoid arthritis, and hyperproliferative disorders.

XX Claim 11; Fig 1; 291pp; English.

XX The invention relates to human and murine peptidoglycan recognition protein-related liver proteins (PGRP-L) and nucleic acid molecules encoding them. The polypeptides of the invention can be used as food additive or preservative to increase or decrease storage capabilities. The PGRP-L polynucleotides are used for chromosome identification. They are also useful as probes for diagnosing disorders related to the female reproductive system, particularly breast and ovarian cancer. They are also useful in the gene therapy of breast and ovarian cancer. The PGRP-L polynucleotides, polypeptides, and their antibodies, agonists and antagonists are useful in the diagnosis, treatment and prevention of cancer particularly breast and ovarian cancer, and cancers of the adrenal gland, gastrointestinal tract, liver, lung, or urogenital; immune disorders such as Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular diseases such as myocardial ischaemias; wound healing; neurological diseases such as Alzheimer's disease, cerebral anoxia and epilepsy; and infectious diseases such as viral, bacterial, fungal and parasitic infections. The present sequence is human peptidoglycan recognition protein- related liver protein (hPGRP-L). This protein has molecular weight of about 18,595 Da

XX SQ Sequence 174 AA;

Query Match 100.0%; Score 174; DB 4; Length 174;  
Best Local Similarity 100.0%; Pred. No. 3.6e-167;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	RGWHWVG	AHTLGHNSRGF	VAIVGN	YTAALPTEAALRTVRD	TLPSCAVRAGLLRPDYALL	60
Db	1	RGWHWVG	AHTLGHNSRGF	VAIVGN	YTAALPTEAALRTVRD	TLPSCAVRAGLLRPDYALL	60
QY	61	GHRQLV	RTDCPGDALF	DLRTWP	HFTAVSLRSLHYTARRP	SVYTSSTRPLPPACNSCART	120
Db	61	GHRQLV	RTDCPGDALF	DLRTWP	HFTAVSLRSLHYTARRP	SVYTSSTRPLPPACNSCART	120
QY	121	ASARPPT	SRRHVYSGNL	GPAGH	SAGNIPDPVTSYAASA	AQPTQACFPFSS	174
Db	121	ASARPPT	SRRHVYSGNL	GPAGH	SAGNIPDPVTSYAASA	AQPTQACFPFSS	174



RESULT 2  
AAG65915  
ID AAG65915 standard; protein; 634 AA.  
XX  
AC AAG65915;  
XX  
DT 11-FEB-2002 (first entry)  
XX  
DE Amino acid sequence of GSK gene Id 239881.  
XX  
KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;  
KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;  
KW cytotstatic; cerebroprotective; vasotropic; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200172961-A2.  
XX  
PD 04-OCT-2001.  
XX  
PF 22-MAR-2001; 2001WO-US009226.  
XX  
PR 24-MAR-2000; 2000US-0192158P.  
PR 28-MAR-2000; 2000US-0192668P.  
PR 27-APR-2000; 2000US-0200166P.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y;  
XX  
DR WPI; 2001-639223/73.  
DR N-PSDB; AAI67205.  
XX  
XX Isolated polypeptides, which may be peptide hormones, which are  
PT identified by high throughput genome-based biology which identifies genes  
PT and gene products as therapeutic targets for treatment of diseases such  
PT as diabetes and cancer.  
XX  
PS Claim 1; Page 94-95; 99pp; English.  
XX  
CC The invention provides polypeptides (AAG65886-65918) which may be peptide  
CC hormones (including insulin, growth hormones, chemokines, cytokines,  
CC neuropeptides, integrins, kallikreins, lamins, melanins, natruliretic  
CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
CC secretogranins, selectins, thromboglobulins, thymosins) identified by  
CC high throughput genome-based biology and polynucleotides (AAI67176-67208)  
CC encoding them. The polypeptides can be expressed by standard recombinant  
CC methodology. The polypeptides are useful in the treatment of disease such  
CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
CC asthma, manic depression, dementia, delirium, mental retardation,  
CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental  
CC or sexual development disorders, and dysfunctions of the blood cascade  
CC system including those leading to stroke. The polynucleotides may be used  
CC as diagnostic reagents through detecting mutations in the associated gene  
CC and for chromosome localization and for tissue expression studies. The  
CC polypeptides and polynucleotides may also be used as vaccines  
XX  
SQ Sequence 634 AA;

Query Match 100.0%; Score 174; DB 4; Length 634;  
Best Local Similarity 100.0%; Pred. No. 1.1e-166;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60  
DB 461 RGWHVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520

QY 61 GHRQLVRTDCPGDALFLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPACNSCART 120  
DB 521 GHRQLVRTDCPGDALFLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPACNSCART 580

QY 121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAASAQPTQACFPFPSS 174  
DB 581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAASAQPTQACFPFPSS 634

RESULT 3  
ADA54695  
ID ADA54695 standard; protein; 634 AA.  
XX  
AC ADA54695;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human protein, SEQ ID 2263.  
XX  
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
KW inflammatory disease; osteoporosis; neurological disease.  
XX  
OS Homo sapiens.  
XX  
PN EPI299569-A2.  
XX  
PD 19-MAR-2003.  
XX  
PF 21-MAR-2002; 2002EP-00006586.  
XX  
PR 14-SEP-2001; 2001JP-00328381.  
PR 24-JAN-2002; 2002US-0350435P.  
XX  
PA (HELI-) HELIX RES INST.  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
DR WPI; 2003-395539/38.  
DR N-PSDB; ADA53056.  
XX  
PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.  
XX  
PS Claim 14; SEQ ID NO 2263; 205pp; English.  
XX  
CC The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.  
XX  
SQ Sequence 634 AA;

Query Match 100.0%; Score 174; DB 6; Length 634;  
Best Local Similarity 100.0%; Pred. No. 1.1e-166;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60  
DB 461 RGWHVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520

QY 61 GHRQLVRTDCPGDALFLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPACNSCART 120  
DB 521 GHRQLVRTDCPGDALFLLRTWPHTAVSLRSLHYTARRPSVYTSSTRPLPACNSCART 580

QY 121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAASAQPTQACFPFPSS 174  
DB 581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAAASAQPTQACFPFPSS 634

RESULT 4

AAAY94863  
ID AAY94863 standard; protein; 363 AA.  
XX  
AC AAY94863;  
XX  
DT 12-JUN-2000 (first entry)  
XX  
DE Human protein clone HP10477.  
XX  
KW Human protein; hydrophobic domain; nutritional source; haematopoiesis;  
KW cytokine production; cell proliferation; cell differentiation;  
KW immune deficiency; infectious disease; autoimmune disorder; asthma;  
KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;  
KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;  
KW nervous system disorder; Alzheimer's disease; Parkinson's disease;  
KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;  
KW systemic cytokine damage; tissue differentiation; contraceptive; stroke;  
KW coagulation disorder; myocardial infarction; inflammatory condition;  
KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;  
KW nephritis; therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200005367-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 22-JUL-1999; 99WO-JP003929.  
XX  
PR 24-JUL-1998; 98JP-00208820.  
PR 07-AUG-1998; 98JP-00224105.  
PR 25-AUG-1998; 98JP-00238116.  
PR 09-SEP-1998; 98JP-00254736.  
PR 29-SEP-1998; 98JP-00275505.  
XX  
PA (SAGA ) SAGAMI CHEM RES CENT.  
PA (PROT-) PROTEGENE INC.  
XX  
PI Kato S, Kimura T;  
XX  
WPI; 2000-182694/16.  
XX  
PT Novel human proteins having hydrophobic domains useful for treating  
PT osteoporosis, Alzheimer's disease, Parkinson's disease, asthma, multiple  
PT sclerosis, rheumatoid arthritis, cancer, anemia, and stroke.  
XX  
PS Claim 1; Page 211-212; 351pp; English.  
XX  
CC This sequence represents a human protein of the invention, which has  
CC hydrophobic domains. The DNA sequences can be used as a probe or as a  
CC genetic marker. The protein can also be used as a marker, and to identify  
CC potential genetic disorders. The DNA and protein can also be used as  
CC nutritional sources or supplements. The protein exhibits cytokine, cell  
CC proliferation, cell differentiation activities and induces production of  
CC other cytokines in certain cell populations. The protein also exhibits  
CC immune stimulating or immune suppressing activity. It can be used in the  
CC treatment of various immune deficiencies and disorders, and to treat  
CC infectious diseases caused by viral, bacterial, fungal or other  
CC infections. The protein is also used for treating autoimmune disorders  
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid  
CC arthritis. It is also useful in the treatment of allergic reactions and  
CC conditions such as asthma, and in immune suppression after organ  
CC transplantation. The protein is useful in regulation of haematopoiesis  
CC and consequently in the treatment of myeloid or lymphoid cell  
CC deficiencies. It is also used in compositions for tissue growth or  
CC regeneration. The protein is also used in the treatment of osteoporosis  
CC or osteoarthritis and in the treatment of periodontal disease and other  
CC tooth repair processes. The protein is used in the treatment of nervous  
CC system disorders such as Alzheimer's disease, Parkinson's disease, and  
CC Huntington's disease. They are useful for protection or regeneration and  
CC treatment of lung or liver fibrosis, reperfusion injury in various  
CC tissues, and conditions resulting from systemic cytokine damage. They are  
CC also used for promoting or inhibiting tissue differentiation. They are

CC also used as contraceptives since they exhibit activin or inhibin related  
CC activities and as a fertility inducing therapeutic. They are used for  
CC treating various coagulation disorders and in treatment and prevention of  
CC conditions resulting from coagulation activities e.g. myocardial  
CC infarction or stroke. They also acts as receptors, receptor ligands or  
CC inhibitors or agonists of receptor/ligand interactions. They are used to  
CC treat inflammatory conditions such as septic shock, sepsis, ischaemia  
CC reperfusion injury, arthritis, and nephritis. They can be used to prevent  
CC tumours  
XX  
SQ Sequence 363 AA;  
Query Match 50.0%; Score 87; DB 3; Length 363;  
Best Local Similarity 100.0%; Pred. No. 4.6e-79;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGHWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60  
Db 248 RGHWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 307  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHFTA 87  
Db 308 GHRQLVRTDCPGDALFDLLRTWPHFTA 334  
RESULT 5  
AAG65916  
ID AAG65916 standard; protein; 576 AA.  
XX  
AC AAG65916;  
XX  
DT 11-FEB-2002 (first entry)  
XX  
DE Amino acid sequence of GSK gene Id 239881.  
XX  
KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;  
KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;  
KW cyostatic; cerebroprotective; vasotropic; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200172961-A2.  
XX  
PD 04-OCT-2001.  
XX  
PF 22-MAR-2001; 2001WO-US009226.  
XX  
PR 24-MAR-2000; 2000US-0192158P.  
PR 28-MAR-2000; 2000US-0192668P.  
PR 27-APR-2000; 2000US-0200166P.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX  
PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;  
PI Lai Y;  
XX  
DR WPI; 2001-639223/73.  
DR N-PSDB; AAI67206.  
XX  
PT Isolated polypeptides, which may be peptide hormones, which are  
PT identified by high throughput genome-based biology which identifies genes  
PT and gene products as therapeutic targets for treatment of diseases such  
PT as diabetes and cancer.  
XX  
PS Claim 1; Page 95-96; 99pp; English.  
XX  
CC The invention provides polypeptides (AAG65886-65918) which may be peptide  
CC hormones (including insulin, growth hormones, chemokines, cytokines,  
CC neuropeptides, integrins, kallikreins, lamins, melanins, natruiretic  
CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,  
CC secretogranins, selectins, thromboglobulins, thymosins) identified by  
CC high throughput genome-based biology and polynucleotides (AAI67176-67208)



CC encoding them. The polypeptides can be expressed by standard recombinant  
CC methodology. The polypeptides are useful in the treatment of disease such  
CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,  
CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,  
CC asthma, manic depression, dementia, delirium, mental retardation,  
CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental  
CC or sexual development disorders, and dysfunctions of the blood cascade  
CC system including those leading to stroke. The polynucleotides may be used  
CC as diagnostic reagents through detecting mutations in the associated gene  
CC and for chromosome localization and for tissue expression studies. The  
CC polypeptides and polynucleotides may also be used as vaccines  
XX  
SQ Sequence 576 AA;

Query Match 50.0%; Score 87; DB 4; Length 576;  
Best Local Similarity 100.0%; Pred. No. 6.9e-79;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHWVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 60  
DB 461 RGWHWVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 520  
QY 61 GHRQLVVRTDCPGDALFDLLRTWPHFTA 87  
DB 521 GHRQLVVRTDCPGDALFDLLRTWPHFTA 547

RESULT 6

ABB78298  
ID ABB78298 standard; protein; 576 AA.

XX ABB78298;

DT 05-DEC-2002 (first entry)

DE Amino acid sequence of lipid-associated molecule (LIPAM)-9.

XX Human; lipid-associated molecule; LIPAM; cardiovascular disease;  
KW atherosclerosis; hypertension; aneurysm; congestive heart failure;  
KW angina pectoris; heart disease; lung disease; oedema; emphysema;  
KW bronchitis; gastrointestinal disease; nausea; peptic ulcer;  
KW Crohn's disease; lipid metabolism; Fabry's disease; diabetes mellitus;  
KW hyperlipidaemia; autoimmune disease; inflammatory disease;  
KW acquired immunodeficiency syndrome; AIDS; anaemia; asthma; gout;  
KW pancreatitis; neurological disease; stroke; Alzheimer's disease;  
KW multiple sclerosis; Parkinson's disease; anxiety; schizophrenia; amnesia;  
KW metabolic disease; Addison's disease; developmental disease;  
KW Cushing's syndrome; endocrine disease; cell proliferative disorder;  
KW cancer; leukemia; lymphoma; sarcoma.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Modified-site 77  
FT /note= "potential glycosylation site"  
FT Modified-site 79  
FT /note= "potential phosphorylation site"  
FT Modified-site 154  
FT /note= "potential phosphorylation site"  
FT Modified-site 181  
FT /note= "potential phosphorylation site"  
FT Modified-site 202  
FT /note= "potential phosphorylation site"  
FT Modified-site 213  
FT /note= "potential phosphorylation site"  
FT Domain 215..234  
FT /note= "transmembrane domain"  
FT Modified-site 239  
FT /note= "potential phosphorylation site"  
FT Domain 255..283  
FT /note= "transmembrane domain"  
FT Modified-site 259  
FT /note= "potential phosphorylation site"

FT Modified-site 367  
FT /note= "potential glycosylation site"  
FT Modified-site 485  
FT /note= "potential glycosylation site"  
FT Modified-site 498  
FT /note= "potential phosphorylation site"  
FT Modified-site 548  
FT /note= "potential phosphorylation site"  
FT Modified-site 558  
FT /note= "potential phosphorylation site"  
FT Modified-site 561  
FT /note= "potential phosphorylation site"

XX WO200263005-A2  
PN 15-AUG-2002.

PF 06-FEB-2002; 2002WO-US003813.  
XX  
PR 06-FEB-2001; 2001US-0266910P.  
PR 16-MAR-2001; 2001US-0276855P.  
PR 16-MAR-2001; 2001US-0276891P.  
PR 28-MAR-2001; 2001US-0279760P.  
PR 13-APR-2001; 2001US-0283818P.  
PR 20-APR-2001; 2001US-0285405P.  
XX

PA (INCY-) INCYTE GENOMICS INC.

XX Das D, Yao MG, Arvizu C, Baughn MR, Lu Y, Hafalia AJA, Walia NK;  
PI Griffin JA, Lu DAM, Yue H, Ding L, Townley DJ, Elliott VS;  
PI Forsythe I, Ramkumar J, Gandhi AR, Ison CH, Warren BA, Tang YT;  
PI Emerling BM, Honchell CD;

DR WPI; 2002-627558/67.  
DR N-PSDB; ABV72135.

XX New human lipid-associated molecules, useful for diagnosing, treating or  
PT preventing cardiovascular disorders (e.g. aneurysms), neurological  
PT disorders (e.g. Parkinson's disease) or cancers (e.g. leukemia or  
PT lymphoma).

PS Claim 1; Page 141-142; 152pp; English.

XX The present sequence represents a lipid-associated molecule (LIPAM).  
CC LIPAM polypeptides, polynucleotides and agonists are useful for treating  
CC a condition associated with decreased expression of functional LIPAM. The  
CC antagonist is useful for treating a disease associated with  
CC overexpression of functional LIPAM. The anti-LIPAM antibody is useful for  
CC diagnosing a condition or disease associated with the expression of  
CC LIPAM. The polypeptides, polynucleotides, agonists and antagonists may  
CC also be used for preventing these diseases. These polypeptides  
CC polynucleotides, agonists and antagonists are particularly useful for  
CC diagnosing, treating or preventing cardiovascular (e.g. atherosclerosis,  
CC hypertension, aneurysms, congestive heart failure, angina pectoris, or  
CC ischaemic or rheumatic heart disease), lung (e.g. oedema, emphysema or  
CC bronchitis) gastrointestinal (e.g. nausea, peptic ulcer or Crohn's  
CC disease), lipid metabolism (e.g. Fabry's disease, diabetes mellitus or  
CC hyperlipidaemia), autoimmune/inflammatory (e.g. acquired immunodeficiency  
CC syndrome (AIDS), anaemia, asthma, gout, pancreatitis or Crohn's disease),  
CC neurological (e.g. stroke, Alzheimer's disease, multiple sclerosis,  
CC Parkinson's disease, anxiety, schizophrenia or amnesia), metabolic (e.g.  
CC Addison's disease), developmental (e.g. Cushing's syndrome), endocrine or  
CC cell proliferative disorders (e.g. cancers including leukemia, lymphoma  
CC or sarcoma)

XX Sequence 576 AA;

Query Match 50.0%; Score 87; DB 5; Length 576;  
Best Local Similarity 100.0%; Pred. No. 6.9e-79;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHWVGAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 60  
|||||

Db 461 RGWHWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 520

QY 61 GHRQLVRTDCPGDALFDLLRTWPHTA 87  
|||||  
521 GHRQLVRTDCPGDALFDLLRTWPHTA 547

Db 521 GHRQLVRTDCPGDALFDLLRTWPHTA 547

RESULT 7  
ABR57566  
ID ABR57566 standard; protein; 576 AA.  
XX  
AC ABR57566;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human Peptidoglycan Recognition Protein, PGRP-L.  
XX  
KW Immunomodulator; gene therapy; human; antimicrobial; chromosome 19;  
KW Peptidoglycan Recognition Protein; PGRP; PGRP-Long; PGRP-L;  
KW PGRP-intermediate alpha; PGRP-intermediate beta; PGRP-Ialpha; PGRP-Ibeta.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Peptide Location/Qualifiers  
FT 1. .21  
FT /label= Signal\_peptide  
FT Protein 22. .576  
FT /label= Mature\_protein  
FT Domain 214. .232  
FT /label= Transmembrane\_domain\_#1  
FT Domain 325. .343  
FT /label= Transmembrane\_domain\_#2  
FT Domain 400. .416  
FT /label= PGRP\_Domain\_III  
FT Domain 442. .470  
FT /label= PGRP\_domain\_II  
FT Domain 495. .545  
FT /label= PGRP\_domain\_I  
XX  
PN WQ2003029401-A2  
XX  
RD 10-APR-2003.  
XX  
PF 15-JUL-2002; 2002WO-US022428.  
XX  
PR 13-JUL-2001; 2001US-0305049P.  
XX  
PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.  
XX  
PI Dziarski R, Liu C, Xu Z, Gupta D;  
XX  
DR WPI; 2003-381614/36.  
DR N-PSDB; ACC70429.  
XX  
PT New peptidoglycan recognition proteins (PGRP)-L, PGRP-I alpha and PGRP-I  
PT beta, useful for modulating the innate immune system, and as research  
PT tools for identifying other proteins involved in regulating antimicrobial  
PT processes.  
XX  
PS Claim 6; Fig 8; 86pp; English.  
XX  
CC The present invention relates to novel human Peptidoglycan Recognition  
CC Proteins (PGRP): PGRP-Long (PGRP-L), PGRP-intermediate alpha and PGRP-  
CC intermediate beta (PGRP-Ialpha and PGRP-Ibeta, ACC70429-ACC70431 and  
CC ABR57566-ABR57568). The PGRPs and their coding sequences are useful for  
CC modulating innate immune system, and as research tools to identify other  
CC proteins that are intimately involved in the regulation of antimicrobial  
CC processes. The gene for PGRP-L is located on chromosome 19 and the genes  
CC for PGRP-Ialpha and PGRP-Ibeta are located on chromosome 1q  
XX  
SQ Sequence 576 AA;

Query Match 50.0%; Score 87; DB 6; Length 576;  
Best Local Similarity 100.0%; Pred. No. 6.9e-79;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 6.9e-79;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 60  
|||||  
461 RGWHWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 520

Db 61 GHRQLVRTDCPGDALFDLLRTWPHTA 87  
|||||  
521 GHRQLVRTDCPGDALFDLLRTWPHTA 547

RESULT 8  
ADE07869  
ID ADE07869 standard; protein; 576 AA.  
XX  
AC ADE07869;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Novel protein (useful for identifying genetic disorders) #24.  
XX  
KW novel gene; novel protein; tissue marker; molecular weight marker;  
KW chromosome marker; genetic disorder.  
XX  
OS Unidentified.  
XX  
PN WQ2003054152-A2.  
XX  
PD 03-JUL-2003.  
XX  
PF 10-DEC-2002; 2002WO-US039555.  
XX  
PR 10-DEC-2001; 2001US-0339739P.  
PR 11-DEC-2001; 2001US-0339453P.  
PR 14-MAR-2002; 2002US-0365091P.  
PR 14-MAR-2002; 2002US-0365384P.  
PR 12-APR-2002; 2002US-0372381P.  
PR 12-APR-2002; 2002US-0372615P.  
PR 22-APR-2002; 2002US-00128558.  
PR 24-APR-2002; 2002US-0376045P.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
XX  
DR WPI; 2003-569235/53.  
DR N-PSDB; ADE06958.  
XX  
PT New polynucleotides, useful for expressing recombinant proteins for  
PT analysis, characterization or therapeutic use, or as markers for tissues  
PT in which the corresponding protein is preferentially expressed.  
XX  
PS Claim 20; SEQ ID NO 935; 1177pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present amino acid sequence represents a protein  
CC of the invention.  
XX  
SQ Sequence 576 AA;

Query Match 50.0%; Score 87; DB 7; Length 576;  
Best Local Similarity 100.0%; Pred. No. 6.9e-79;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHWVGAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 60



```
|||||
461 RGHWVGAAHTLGHNSRGFGVAIVGNYYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 520
61 GHRQLVRTDCPGDALFDLLRTWPHFTA 87
|||||
521 GHRQLVRTDCPGDALFDLLRTWPHFTA 547

RESULT 9
AAY72664
ID AAY72664 standard; protein; 530 AA.
XX
AC AAY72664;
XX
DT 31-MAY-2001 (first entry)
XX
DE Murine peptidoglycan recognition protein-related liver protein.
XX
KW Peptidoglycan recognition protein-related liver protein; PGRP-L;
KW food additive; food preservative; breast cancer; ovarian cancer;
KW immune disorder; Addison's disease; allergy; cardiovascular disorder;
KW myocardial ischaemia; wound healing; neurological disease; vasotropic;
KW Alzheimer's disease; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cerebroprotective; antibacterial;
KW virucide; fungicide; ophthalmological; murine; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .22
FT /label= Signal_peptide
FT 18. .24
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 23. .530
FT /label= Mature_murine_peptidoglycan_recognition_ protein-
FT related_liver_protein
FT 37. .45
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 59. .68
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 76. .93
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 101. .111
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 116. .122
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 143. .159
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 172. .197
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 213. .229
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 242. .246
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 268. .275
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 288. .305
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
FT 325. .330
FT /note= "Antigenic region of murine peptidoglycan
FT recognition protein-related liver protein"
```

```
335. .340
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
348. .353
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
362. .382
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
397. .399
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
406. .413
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
415. .426
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
432. .435
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
440. .442
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
449. .457
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
466. .475
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
478. .484
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
486. .499
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
511. .513
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
521. .530
/note= "Antigenic region of murine peptidoglycan
recognition protein-related liver protein"
XX
PN WC266414545-A1.
XX PD 01-MAR-2001.
XX PF 18-AUG-2000; 2000WO-US022877.
XX PR 20-AUG-1999; 99US-0149715P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Young PE, Rosen CA, Duan RD;
XX WPI; 2001-160115/16.
DR N-PSDB; AAD02743.
XX
PT Isolated nucleic acids encoding human and murine peptidoglycan
PT recognition protein-related liver (PGRP-L) proteins, useful for
PT preventing or treating autoimmune diseases e.g. rheumatoid arthritis, and
PT hyperproliferative disorders.
XX
PS Disclosure; Fig 2; 291pp; English.
XX
CC The invention relates to human and murine peptidoglycan recognition
CC protein-related liver proteins (PGRP-L) and nucleic acid molecules
CC encoding them. The polypeptides of the invention can be used as food
CC additive or preservative to increase or decrease storage capabilities.
CC The PGRP-L polynucleotides are used for chromosome identification. They
CC are also useful as probes for diagnosing disorders related to the female
CC reproductive system, particularly breast and ovarian cancer. They are also
CC useful in the gene therapy of breast and ovarian cancer. The PGRP-L
```

CC polynucleotides, polypeptides, and their antibodies, agonists and  
CC antagonists are useful in the diagnosis, treatment and prevention of  
CC cancer particularly breast and ovarian cancer, and cancers of the adrenal  
CC gland, gastrointestinal tract, liver, lung, or urogenital; immune  
CC disorders such as Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischaemias; wound healing;  
CC neurological diseases such as Alzheimer's disease, cerebral anoxia and  
CC epilepsy; and infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections. The present sequence is murine peptidoglycan  
CC recognition protein- related liver protein (mPGRP-L). This protein has  
CC molecular weight of about 57,764 Da  
XX  
SQ Sequence 530 AA;

Query Match 5.7%; Score 10; DB 4; Length 530;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHWVGAHT 10  
Db 441 RGWHWVGAHT 450

RESULT 10  
AAU40317  
ID AAU40317 standard; protein; 107 AA.  
XX  
AC AAU40317;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #1213.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US012865.  
XX  
PR 21-APR-2000; 2000US-0199047P.  
PR 02-JUN-2000; 2000US-0208841P.  
PR 07-JUL-2000; 2000US-0216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI; 2001-616774/71.  
DR N-PSDB; AAS59511.  
XX  
PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.  
XX  
PS Example 1; SEQ ID NO 1512; 1069pp; English.  
XX  
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 107 AA;

Query Match 4.6%; Score 8; DB 4; Length 107;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 RGFGVAIV 23  
Db 17 RGFGVAIV 24

RESULT 11  
ABM36836  
ID ABM36836 standard; protein; 107 AA.  
XX  
AC ABM36836;  
XX  
DT 20-OCT-2003 (first entry)  
XX  
DE Propionibacterium acnes predicted ORF-encoded polypeptide #1512.  
XX  
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;  
KW immunostimulant; immune response; vaccine.  
XX  
OS Propionibacterium acnes.  
XX  
PN WO2003033515-A1.  
XX  
PD 24-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032727.  
XX  
PR 15-OCT-2001; 2001US-00978825.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
PI Barth B, Vallieve-Douglas J;  
XX  
DR WPI; 2003-381789/36.  
DR N-PSDB; ACF64440.  
XX  
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
PT or for stimulating an immune response specific for a P. acnes protein.  
XX  
PS Example 1; SEQ ID NO 1512; 1481pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,



CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 107 AA;

Query Match 4.6%; Score 8; DB 6; Length 107;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 RGFQVAIV 23  
| | | | |  
Db 17 RGFQVAIV 24

RESULT 12  
ABG17167  
ID ABG17167 standard; protein; 126 AA.  
XX  
AC ABG17167;  
XX

DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #17158.  
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX

OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS81354.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX PS Claim 20; SEQ ID NO 47526; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 126 AA;

Query Match 4.6%; Score 8; DB 4; Length 126;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 CAVRAGLL 53  
| | | | |  
Db 58 CAVRAGLL 65

RESULT 13  
ABB63392  
ID ABB63392 standard; protein; 173 AA.  
XX  
AC ABB63392;  
XX

DT 26-MAR-2002 (first entry)  
XX

DE Drosophila melanogaster polypeptide SEQ ID NO 16968.

KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL07495.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.

XX PS Disclosure; SEQ ID NO 16968; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly

```
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 173 AA;
SQ
    Query Match      4.6%; Score 8; DB 4; Length 173;
    Best Local Similarity 100.0%; Pred. No. 13;
    Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 123 ARPPTSRR 130
Db 102 ARPPTSRR 109
RESULT 14
AAG35998
ID AAG35998 standard; protein; 179 AA.
XX
AC AAG35998;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 44055.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN BP10333405-AA
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
```



PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 4.6%; Score 8; DB 3; Length 179;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ARPPTSRR 130  
Db |||||  
20 ARPPTSRR 27

RESULT 15

AAAY00051  
ID AAAY00051 standard; protein; 526 AA.

XX  
AC AAAY00051;

XX  
DT 20-APR-1999 (first entry)

XX Enterococcus faecalis antigenic polypeptide fragment EF022.

KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;  
XX detection; attenuation; antigenic.

OS Enterococcus faecalis.

XX  
PN WO9850554-A2.

XX  
PD 12-NOV-1998.

XX  
PF 04-MAY-1998; 98WO-US008959.

XX  
PR 06-MAY-1997; 97US-0044031P.

PR 16-MAY-1997; 97US-0046655P.

PR 14-NOV-1997; 97US-00666009P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX  
PI Kunsch CA, Choi GH, Bailey C, Hromockyj A;

DR WPI; 1999-070095/06.

DR N-PSDB; AAX20041.

XX  
PT New isolated Enterococcus faecalis polynucleotides - used to develop  
PT products for the detection of Enterococcus and for use in vaccines for  
PT prevention or attenuation of Enterococcus infection.

XX  
PS Claim 9; Page 102-103; 301pp; English.

XX  
CC The present sequence represents an antigenic polypeptide fragment  
CC isolated from Enterococcus faecalis. The present invention describes  
CC genes, proteins and antigenic polypeptides isolated from E. faecalis. The  
CC proteins can be used in vaccines for preventing or attenuating an  
CC infection caused by a member of the Enterococcus genus in an animal. They  
CC can also be used for detecting Enterococcus antibodies in a sample. The  
CC nucleotide sequences can be used for detecting Enterococcus nucleic  
CC acids. Products from the present invention can also be used for screening  
CC compounds to identify agonists and antagonists of E. faecalis protein  
CC activity

SQ Sequence 526 AA;

Query Match 4.6%; Score 8; DB 2; Length 526;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ALPTEAAL 36

Db |||||  
404 ALPTEAAL 411

RESULT 16

ABP43270  
ID ABP43270 standard; protein; 526 AA.

XX

AC ABP43270;

XX

DT 05-AUG-2002 (first entry)

XX DE E faecalis EF022 antigenic fragment.  
XX KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.  
XX OS Enterococcus faecalis.  
XX PN US2002045737-A1.  
XX PD 18-APR-2002.  
XX PF 04-MAY-1998; 98US-00071035.  
XX PR 04-MAY-1998; 98US-00071035.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
XX DR WPI; 2002-425450/45.  
XX DR N-PSDB; ABN98026.  
XX PT New genes and polypeptides from Enterococcus faecalis, useful as vaccines  
PT for preventing, treating or attenuating an infection caused by a member  
PT of the Enterococcus genus in an animal, particularly E. faecalis.  
XX PS Claim 9; Page 55-56; 255pp; English.  
XX CC The present invention provides the protein and coding sequences of a  
CC number of polypeptides from Enterococcus faecalis. The proteins can be  
CC used as vaccines for preventing or attenuating an infection caused by a  
CC member of the Enterococcus genus in an animal, particularly E. faecalis.  
CC The polynucleotide is also useful for preventing or treating E. faecalis  
CC infection. The present sequence is a protein of the invention  
XX SQ Sequence 526 AA;  
Query Match 4.6%; Score 8; DB 5; Length 526;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 ALPTEAAL 36  
Db |||||  
404 ALPTEAAL 411  
RESULT 17  
ABU88298  
ID ABU88298 standard; protein; 526 AA.  
XX AC ABU88298;  
XX DT 07-JUL-2003 (first entry)  
XX DE E. faecalis novel protein #42.  
XX KW Vaccine; endocarditis; bacteraemia; urinary tract infection; UTI;  
KW intraabdominal infection; soft tissue infection; neonatal sepsis.  
XX OS Enterococcus faecalis.  
XX PN US2003017495-A1.  
XX PD 23-JAN-2003.  
XX PF 29-JUL-2002; 2002US-00206576.  
XX PR 06-MAY-1997; 97US-0044031P.  
PR 16-MAY-1997; 97US-0046655P.  
PR 14-NOV-1997; 97US-0066009P.  
PR 04-MAY-1998; 98US-00071035.  
XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
XX DR WPI; 2003-416890/39.  
DR N-PSDB; ACA87990.  
XX PT New nucleic acid molecules and polypeptides from Enterococcus faecalis,  
PT useful as vaccines for preventing or attenuating an enterococcal  
PT infection in an animal, or for identifying Enterococcus faecalis in  
PT biological samples.  
XX PS Claim 12; Page; 40pp; English.  
XX CC The invention relates to a new isolated nucleic acid molecule comprising  
CC a polynucleotide isolated from Enterococcus faecalis appearing as  
CC ACA87949-ACA88196 (or sequences complementary to them or 95% identical to  
CC them). Also included are the proteins encoded by the above nucleic acids,  
CC making a recombinant vector (comprising inserting the isolated nucleic  
CC acid molecule cited above into a vector), a host cell comprising the  
CC vector (used to produce the protein), an isolated antibody specific for  
CC the polypeptides, a hybridoma that produces the antibody, an isolated  
CC polypeptide antigen comprising an amino acid sequence of an Enterococcus  
CC faecalis epitope listed in the specification, a vaccine comprising one or  
CC more E. faecalis polypeptides (and a pharmaceutical diluent, carrier or  
CC excipient) where the polypeptide elicits protective antibodies in an  
CC animal to a member of the genus Enterococcus; preventing or attenuating  
CC an infection caused by a member of the genus Enterococcus in an animal  
CC comprising administering to the animal the polypeptide and detecting  
CC Enterococcus nucleic acids in a biological sample. The E. faecalis  
CC nucleic acid molecules and polypeptides are useful as vaccines for  
CC preventing or attenuating an enterococcal infection in an animal (e.g.  
CC endocarditis, bacteraemia, urinary tract infection (UTI), intraabdominal  
CC infection, soft tissue infection and neonatal sepsis). The polypeptides  
CC are also useful for detecting Enterococcus aureus in immunoassays, as  
CC epitope tags, as molecular weight markers, or for generating antibodies  
CC that specifically bind E. faecalis polypeptides. The nucleic acid  
CC molecules are also useful as probes for gene mapping, or for identifying  
CC E. faecalis in biological samples. The kit and methods are useful for  
CC detecting Enterococcus antibodies or nucleic acid molecules in a  
CC biological sample. The present sequence is a novel E. faecalis  
CC polypeptide of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from the USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030017495  
XX SQ Sequence 526 AA;  
Query Match 4.6%; Score 8; DB 6; Length 526;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 ALPTEAAL 36  
Db |||||  
404 ALPTEAAL 411  
RESULT 18  
ABU13549  
ID ABU13549 standard; protein; 526 AA.  
XX AC ABU13549;  
XX DT 26-FEB-2003 (first entry)  
XX DE Enterococcus faecalis EF040 polypeptide #42.  
XX KW EF040; immunostimulant; antibacterial; gene mapping.  
XX OS Enterococcus faecalis.  
XX PN US6448043-B1.  
XX PD 10-SEP-2002.



XX 04-MAY-1998; 98US-00071035.  
PF 06-MAY-1997; 97US-0044031P.  
XX 16-MAY-1997; 97US-0046655P.  
PR 14-NOV-1997; 97US-0066009P.  
XX 14-NOV-1997; 97US-0066099P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
XX  
DR WPI; 2003-089120/08.  
DR N-PSDB; ABX61596.  
XX  
PT New EF040 polypeptides and polynucleotides from Enterococcus faecalis,  
PT useful for generating an immune response against E. faecalis and other  
PT Enterococcus species, and as vaccines against other bacterial genera.  
XX  
PS Example 1; Col 75-78; 146pp; English.  
XX  
CC The invention relates to polynucleotide fragments of a gene from  
CC Enterococcus faecalis, EF040, and the polypeptides encoded by them. The  
CC polypeptides are useful in detecting E. faecalis, as epitope tags, as  
CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel  
CC filtration columns, in generating antibodies that specifically bind to  
CC the E. faecalis polypeptides, in generating an immune response against E.  
CC faecalis and other Enterococcus species and as vaccines against other  
CC bacterial genera. The polynucleotides are useful as probes for gene  
CC mapping and for identifying E. faecalis in biological samples. Sequences  
CC ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:  
CC The sequence data for this patent can also be obtained from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 526 AA;  
  
Query Match 4.6%; Score 8; DB 6; Length 526;  
Best Local Similarity 100.0%; Pred.No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 29 ALPTEAAL 36  
Db 404 ALPTEAAL 411  
|||||  
  
RESULT 19  
AAY00050  
ID AAY00050 standard; protein; 546 AA.  
XX  
AC AAY00050;  
XX  
DT 20-APR-1999 (first entry)  
XX  
DE Enterococcus faecalis protein EF022.  
XX  
KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;  
KW detection; attenuation; antigenic.  
XX  
OS Enterococcus faecalis.  
XX  
PN WO9850554-A2.  
XX  
PD 12-NOV-1998.  
XX  
PF 04-MAY-1998; 98WO-US008959.  
XX  
PR 06-MAY-1997; 97US-0044031P.  
PR 16-MAY-1997; 97US-0046655P.  
PR 14-NOV-1997; 97US-0066009P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Kunsch CA, Choi GH, Bailey C, Hromockyj A;

XX WPI; 1999-070095/06.  
DR N-PSDB; AAX20040.  
XX  
PT New isolated Enterococcus faecalis polynucleotides - used to develop  
PT products for the detection of Enterococcus and for use in vaccines for  
PT prevention or attenuation of Enterococcus infection.  
XX  
PS Claim 9; Page 102; 301pp; English.  
XX  
CC The present sequence represents a protein isolated from Enterococcus  
CC faecalis. The present invention describes genes, proteins and antigenic  
CC polypeptides isolated from E. faecalis. The proteins can be used in  
CC vaccines for preventing or attenuating an infection caused by a member of  
CC the Enterococcus genus in an animal. They can also be used for detecting  
CC Enterococcus antibodies in a sample. The nucleotide sequences can be used  
CC for detecting Enterococcus nucleic acids. Products from the present  
CC invention can also be used for screening compounds to identify agonists  
CC and antagonists of E. faecalis protein activity  
XX  
SQ Sequence 546 AA;  
  
Query Match 4.6%; Score 8; DB 2; Length 546;  
Best Local Similarity 100.0%; Pred.No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 29 ALPTEAAL 36  
Db 424 ALPTEAAL 431  
|||||  
  
RESULT 20  
ABP43269  
ID ABP43269 standard; protein; 546 AA.  
XX  
AC ABP43269;  
XX  
DT 05-AUG-2002 (first entry)  
XX  
DE E faecalis EF022 protein.  
XX  
KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.  
XX  
OS Enterococcus faecalis.  
XX  
PN US2002045737-A1.  
XX  
PD 18-APR-2002.  
XX  
PF 04-MAY-1998; 98US-00071035.  
XX  
PR 04-MAY-1998; 98US-00071035.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;  
XX  
DR WPI; 2002-425450/45.  
DR N-PSDB; ABN98025.  
XX  
PT New genes and polypeptides from Enterococcus faecalis, useful as vaccines  
PT for preventing, treating or attenuating an infection caused by a member  
PT of the Enterococcus genus in an animal, particularly E. faecalis.  
XX  
PS Claim 9; Page 54-55; 255pp; English.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of polypeptides from Enterococcus faecalis. The proteins can be  
CC used as vaccines for preventing or attenuating an infection caused by a  
CC member of the Enterococcus genus in an animal, particularly E. faecalis.  
CC The polynucleotide is also useful for preventing or treating E. faecalis  
CC infection. The present sequence is a protein of the invention  
XX

SQ      Sequence 546 AA;

Query Match                      4.6%;    Score 8;    DB 5;    Length 546;  
Best Local Similarity    100.0%;    Pred. No. 36;  
Matches    8;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

QY                      29 ALPTEAAL 36  
                    |||||

Db                      424 ALPTEAAL 431

RESULT 21  
ABU88297

ID    ABU88297 standard; protein; 546 AA.  
XX  
AC    ABU88297;  
XX  
DT    07-JUL-2003    (first entry)  
XX  
DE    E. faecalis novel protein #41.  
XX  
KW    Vaccine; endocarditis; bacteraemia; urinary tract infection; UTI;  
KW    intraabdominal infection; soft tissue infection; neonatal sepsis.  
XX  
OS    Enterococcus faecalis.  
XX  
PN    US2003017495-A1.  
XX  
PD    23-JAN-2003.  
XX  
PF    29-JUL-2002; 2002US-00206576.  
XX  
PR    06-MAY-1997;    97US-0044031P.  
PR    16-MAY-1997;    97US-0046655P.  
PR    14-NOV-1997;    97US-0066009P.  
PR    04-MAY-1998;    98US-00071035.  
XX  
PA    (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI    Choi GH,    Bailey C,    Hromockyj A,    Kunsch CA;  
XX  
DR    WPI; 2003-416890/39.  
DR    N-PSDB; ACA87989.  
XX  
PT    New nucleic acid molecules and polypeptides from Enterococcus faecalis,  
PT    useful as vaccines for preventing or attenuating an enterococcal  
PT    infection in an animal, or for identifying Enterococcus faecalis in  
PT    biological samples.  
XX  
PS    Claim 12; Page; 40pp; English.  
XX  
CC    The invention relates to a new isolated nucleic acid molecule comprising  
CC    a polynucleotide isolated from Enterococcus faecalis appearing as  
CC    ACA87949-ACA88196 (or sequences complementary to them or 95% identical to  
CC    them). Also included are the proteins encoded by the above nucleic acids,  
CC    making a recombinant vector (comprising inserting the isolated nucleic  
CC    acid molecule cited above into a vector), a host cell comprising the  
CC    vector (used to produce the protein), an isolated antibody specific for  
CC    the polypeptides, a hybridoma that produces the antibody, an isolated  
CC    polypeptide antigen comprising an amino acid sequence of an Enterococcus  
CC    faecalis epitope listed in the specification, a vaccine comprising one or  
CC    more E. faecalis polypeptides (and a pharmaceutical diluent, carrier or  
CC    excipient) where the polypeptide elicits protective antibodies in an  
CC    animal to a member of the genus Enterococcus; preventing or attenuating  
CC    an infection caused by a member of the genus Enterococcus in an animal  
CC    comprising administering to the animal the polypeptide and detecting  
CC    Enterococcus nucleic acids in a biological sample. The E. faecalis  
CC    nucleic acid molecules and polypeptides are useful as vaccines for  
CC    preventing or attenuating an enterococcal infection in an animal (e.g.  
CC    endocarditis, bacteraemia, urinary tract infection (UTI), intraabdominal  
CC    infection, soft tissue infection and neonatal sepsis). The polypeptides  
CC    are also useful for detecting Enterococcus aureus in immunoassays, as  
CC    epitope tags, as molecular weight markers, or for generating antibodies

CC    that specifically bind E. faecalis polypeptides. The nucleic acid  
CC    molecules are also useful as probes for gene mapping, or for identifying  
CC    E. faecalis in biological samples. The kit and methods are useful for  
CC    detecting Enterococcus antibodies or nucleic acid molecules in a  
CC    biological sample. The present sequence is a novel E. faecalis  
CC    polypeptide of the invention. Note: The sequence data for this patent did  
CC    not form part of the printed specification, but was obtained in  
CC    electronic format directly from the USPTO at  
CC    seqdata.uspto.gov/sequence.html?DocID=20030017495  
XX  
SQ      Sequence 546 AA;

Query Match                      4.6%;    Score 8;    DB 6;    Length 546;  
Best Local Similarity    100.0%;    Pred. No. 36;  
Matches    8;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

QY                      29 ALPTEAAL 36  
                    |||||

Db                      424 ALPTEAAL 431

RESULT 22  
ABU13548

ID    ABU13548 standard; protein; 546 AA.  
XX  
AC    ABU13548;  
XX  
DT    26-FEB-2003    (first entry)  
XX  
DE    Enterococcus faecalis EF040 polypeptide #41.  
XX  
KW    EF040; immunostimulant; antibacterial; gene mapping.  
XX  
OS    Enterococcus faecalis.  
XX  
PN    US6448043-B1.  
XX  
PD    10-SEP-2002.  
XX  
PF    04-MAY-1998;    98US-00071035.  
XX  
PR    06-MAY-1997;    97US-0044031P.  
PR    16-MAY-1997;    97US-0046655P.  
PR    14-NOV-1997;    97US-0066009P.  
PR    14-NOV-1997;    97US-0066099P.  
XX  
PA    (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI    Choi GH,    Bailey C,    Hromockyj A,    Kunsch CA;  
XX  
DR    WPI; 2003-089120/08.  
DR    N-PSDB; ABX61595.  
XX  
PT    New EF040 polypeptides and polynucleotides from Enterococcus faecalis,  
PT    useful for generating an immune response against E. faecalis and other  
PT    Enterococcus species, and as vaccines against other bacterial genera.  
XX  
PS    Example 1; Col 75-76; 146pp; English.  
XX  
CC    The invention relates to polynucleotide fragments of a gene from  
CC    Enterococcus faecalis, EF040, and the polypeptides encoded by them. The  
CC    polypeptides are useful in detecting E. faecalis, as epitope tags, as  
CC    molecular weight markers on SDS-PAGE gels or for molecular sieve gel  
CC    filtration columns, in generating antibodies that specifically bind to  
CC    the E. faecalis polypeptides, in generating an immune response against E.  
CC    faecalis and other Enterococcus species and as vaccines against other  
CC    bacterial genera. The polynucleotides are useful as probes for gene  
CC    mapping and for identifying E. faecalis in biological samples. Sequences  
CC    ABU13508-ABU13755 represent EF040 polypeptides of the invention. Note:  
CC    The sequence data for this patent can also be obtained from USPTO at  
CC    seqdata.uspto.gov/sequence.html  
XX  
SQ      Sequence 546 AA;



Query Match 4.6%; Score 8; DB 6; Length 546;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ALPTEAAL 36  
| | | | | | | |  
Db 424 ALPTEAAL 431

RESULT 23  
AAAY57445  
ID AAY57445 standard; protein; 1197 AA.  
XX  
AC AAY57445;  
XX  
DT 28-FEB-2000 (first entry)  
XX  
DE Mouse Ese2 protein sequence.  
XX  
KW Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;  
KW regulation; actin cytoskeleton; detection; cancer; infection;  
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;  
KW antiproliferative; antiviral.

OS Mus sp.  
XX  
PN WO9955728-A2.  
XX  
PD 04-NOV-1999.  
XX  
PF 27-APR-1999; 99WO-CA000375.  
XX  
PR 27-APR-1998; 98CA-02230201.  
PR 05-FEB-1999; 99US-0118739P.  
XX  
PA (HSCR-) HSC RES & DEV LP.  
XX  
PI Egan SE, Wang W, Sengar A;  
XX  
DR WPI; 2000-052802/04.  
DR N-PSDB; AAZ39010, AAZ39011.

XX  
PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of  
PT endocytosis, used e.g. for treating cancer or preventing viral infection.  
XX  
PS Claim 33; Page 48; 99pp; English.

XX  
CC The present sequence represents mouse Ese2. The present invention  
CC specifically describes mammalian Ese1 and 2 proteins (I) and their splice  
CC variants (Ese = EH-domain and SH3-domain regulator of endocytosis). (I)  
CC are involved in regulation of clathrin-mediated endocytosis (as a complex  
CC with Esp15 protein), vesicular trafficking and actin cytoskeleton.  
CC Generally (I) (or its (ant)agonists, mimetics, fragments and inactive  
CC mutants); (I)-specific antibodies (Ab); sequences antisense to the (I)  
CC polynucleotide; agents that downregulate expression of Ese genes or  
CC antagonists of an Ese binding partner are used to treat diseases  
CC associated with undesirable endocytosis and resulting changes in cellular  
CC function. Particularly overexpression of Ese1 is used to block clathrin-  
CC mediated endocytosis in vivo or in cell cultures, while administration of  
CC (I) is used to promote endocytosis of selected cells. (Ant)agonists of  
CC (I) or Ab are used to suppress abnormal proliferation of cells that can  
CC be stimulated to proliferate by a growth factor receptor; and similar  
CC compounds (also inactive Ese mutants) can be used to prevent viral  
CC infection. Endocytosis may also be regulated, in vivo or in cell  
CC cultures, by forming an Ese-Esp15 complex, then binding dynamitin to the  
CC complex. Generally conditions that can be treated include cancer;  
CC abnormal cell division or migration; viral infection; or abnormal  
CC receptor signalling, tissue development or synaptic transmission

XX  
SQ Sequence 1197 AA;

Query Match 4.6%; Score 8; DB 3; Length 1197;

Best Local Similarity 100.0%; Pred. No. 73;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 VTSAYAAS 160  
| | | | | | | |  
Db 1020 VTSAYAAS 1027

RESULT 24  
AAAY57450  
ID AAY57450 standard; protein; 1658 AA.  
XX  
AC AAY57450;  
XX  
DT 28-FEB-2000 (first entry)  
XX  
DE Mouse Ese2L protein sequence.  
XX  
KW Mouse; murine; Ese1; Ese2; endocytosis; vesicular trafficking;  
KW regulation; actin cytoskeleton; detection; cancer; infection;  
KW EH-domain and SH3-domain regulator of endocytosis; anticancer;  
KW antiproliferative; antiviral.

OS Mus sp.  
XX  
PN WO9955728-A2.  
XX  
PD 04-NOV-1999.  
XX  
PF 27-APR-1999; 99WO-CA000375.  
XX  
PR 27-APR-1998; 98CA-02230201.  
PR 05-FEB-1999; 99US-0118739P.  
XX  
PA (HSCR-) HSC RES & DEV LP.  
XX  
PI Egan SE, Wang W, Sengar A;  
XX  
DR WPI; 2000-052802/04.  
DR N-PSDB; AAZ39026, AAZ39027.

XX  
PT New nucleic acid encoding Ese1 and 2 proteins, involved in regulation of  
PT endocytosis, used e.g. for treating cancer or preventing viral infection.  
XX  
PS Claim 33; Page 69-70; 99pp; English.

XX  
CC The present invention specifically describes mammalian Ese1 and 2  
CC proteins (I) and their splice variants (Ese = EH-domain and SH3-domain  
CC regulator of endocytosis). (I) are involved in regulation of clathrin-  
CC mediated endocytosis (as a complex with Esp15 protein), vesicular  
CC trafficking and actin cytoskeleton. Generally (I) (or its (ant)agonists,  
CC mimetics, fragments and inactive mutants); (I)-specific antibodies (Ab);  
CC sequences antisense to the (I) polynucleotide; agents that downregulate  
CC expression of Ese genes or antagonists of an Ese binding partner are used  
CC to treat diseases associated with undesirable endocytosis and resulting  
CC changes in cellular function. Particularly overexpression of Ese1 is used  
CC to block clathrin-mediated endocytosis in vivo or in cell cultures, while  
CC administration of (I) is used to promote endocytosis of selected cells.  
CC (Ant)agonists of (I) or Ab are used to suppress abnormal proliferation of  
CC cells that can be stimulated to proliferate by a growth factor receptor;  
CC and similar compounds (also inactive Ese mutants) can be used to prevent  
CC viral infection. Endocytosis may also be regulated, in vivo or in cell  
CC cultures, by forming an Ese-Esp15 complex, then binding dynamitin to the  
CC complex. Generally conditions that can be treated include cancer;  
CC abnormal cell division or migration; viral infection; or abnormal  
CC receptor signalling, tissue development or synaptic transmission. The  
CC present sequence represents mouse Ese2L protein sequence

XX  
SQ Sequence 1658 AA;

Query Match 4.6%; Score 8; DB 3; Length 1658;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 VTSAYAAS 160  
Db 1020 VTSAYAAS 1027

RESULT 25  
AAAY71160  
ID AAAY71160 standard; protein; 1683 AA.  
XX  
AC AAAY71160;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Rat phosphodiesterase interacting protein, M14.  
XX  
KW Rat; phosphodiesterase interacting protein; M14; PDE; cAMP-PDE;  
KW cyclic adenosine monophosphate phosphodiesterase; antiasthmatic; asthma;  
KW antiinflammatory; antipsoriatic; dermatological; antibacterial; shock;  
KW analgesic; immunosuppressive; antiulcer; vasotropic; antiarthritic;  
KW antidiabetic; urticaria; antiallergic; antiarteriosclerotic; diagnosis;  
KW antirheumatic; treatment; inflammatory disease; psoriasis; arthritis;  
KW atherosclerosis; Crohn's disease; cystic fibrosis; chronic bronchitis;  
KW eosinophilic granuloma; proliferative skin disease; ulcerative colitis;  
KW reperfusion injury; atopic dermatitis; diabetes insipidus;  
KW conjunctivitis; adult respiratory distress syndrome; allergic rhinitis;  
KW arterial restenosis; ankylosing spondylitis; transplant rejection;  
KW graft versus host disease.  
XX  
OS Rattus sp.  
XX  
PN WO200027861-A1.  
XX  
PD 18-MAY-2000.  
XX  
PF 12-NOV-1999; 99WO-US026860.  
XX  
PR 12-NOV-1998; 98US-0108255P.  
XX  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
XX  
PI Conti M, Pahlke G;  
XX  
DR WPI; 2000-376479/32.  
XX  
PT Polynucleotide encoding a phosphodiesterase (PDE) interacting  
PT polypeptide, useful for diagnosis and treatment of asthma, cystic  
PT fibrosis, Crohn's disease, and rheumatoid arthritis.  
XX  
PS Disclosure; Fig 6; 77pp; English.  
XX  
CC The present sequence is a phosphodiesterase (PDE) interacting protein,  
CC M14 from rat. The protein modulates the functions and properties of PDEs,  
CC specifically cAMP-PDEs, and also targets them to specific subcellular  
CC compartments. The present sequence can be used in the diagnosis and  
CC treatment of disease conditions associated with PDE activity. The  
CC diseases include asthma, cystic fibrosis, inflammatory airway disease,  
CC chronic bronchitis, eosinophilic granuloma, psoriasis, proliferative skin  
CC diseases, endotoxin shock, septic shock, ulcerative colitis, Crohn's  
CC disease, reperfusion injury, inflammatory arthritis, atopic dermatitis,  
CC urticaria, adult respiratory distress syndrome, diabetes insipidus,  
CC allergic rhinitis, allergic conjunctivitis, vernal conjunctivitis,  
CC arterial restenosis, atherosclerosis, inflammatory diseases associated  
CC with irritation and pain, rheumatoid arthritis, ankylosing spondylitis,  
CC transplant rejection and graft versus host disease, disease conditions  
CC associated with hypersecretion of gastric acid, and disease conditions in  
CC which cytokines are mediators  
XX  
SQ Sequence 1683 AA;

Query Match 4.6%; Score 8; DB 3; Length 1683;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 VTSAYAAS 160  
Db 1046 VTSAYAAS 1053

RESULT 26  
AAR93307  
ID AAR93307 standard; peptide; 10 AA.  
XX  
AC AAR93307;  
XX  
DT 24-APR-1996 (first entry)  
XX  
DE PI3K protein tyrosine kinase derived peptide #1.  
XX  
KW SH3 ligand; SH3 binding agent; biased phage library;  
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;  
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;  
KW p67; complex; chronic myelogenous leukaemia; cancer.  
XX  
OS Synthetic.  
XX  
PN WO9524419-A1.  
XX  
PD 14-SEP-1995.  
XX  
PF 13-MAR-1995; 95WO-US003208.  
XX  
PR 11-MAR-1994; 94US-00209835.  
PR 06-JAN-1995; 95US-00369832.  
XX  
PA (ARIA-) ARIAD PHARM INC.  
XX  
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;  
XX  
DR WPI; 1995-328231/42.  
XX  
PT Identification of peptide(s) binding specifically to SH3 domains - for  
PT use in inhibiting interactions mediated by SH3 domains in treatment of  
PT e.g. osteoporosis and cancer.  
XX  
PS Disclosure; Fig 1; 74pp; English.  
XX  
CC The sequences given in AAR93272-342 represent peptides which are SH3  
CC ligands/SH3 binding agents. They represent a biased phage library which  
CC comprises six random amino acids flanking the tetrapeptide -ppip which  
CC was identified as a recognition sequence for the src SH3 domain. These  
CC sequences were identified using the method of the invention. The method  
CC comprises contacting the SH3 domain with a mixture of peptides under  
CC conditions permitting a ligand to bind to an SH3 domain to form a  
CC complex. Any unbound peptides are removed and the complexed peptide  
CC ligands are dissociated from the complexes. The selected peptides are  
CC enriched by re-contacting them with the SH3 domain and then candidates  
CC which bind to the SH3 domain are detected. The isolated SH3 binding  
CC peptides may be used in the diagnosis, prevention and treatment of  
CC conditions or diseases resulting from cellular processes mediated by an  
CC SH3-based interaction. Such diseases include Paget's disease. Other  
CC conditions treatable with these peptides include restenosis, rheumatoid  
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase  
CC p47 and p67 complex is implicated, etc  
XX  
SQ Sequence 10 AA;

Query Match 4.0%; Score 7; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLPP 112  
Db 2 STRPLPP 8



```
RESULT 27
AAW11112
ID AAW11112 standard; peptide; 13 AA.
XX
AC AAW11112;
XX
DT 25-JUN-1997 (first entry)
XX
DE Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US009382.
XX
PR 22-JUL-1994; 94US-00278865.
PR 07-JUN-1995; 95US-00483555.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
XX
DR WPI; 1996-117151/12.
XX
PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.
XX
PS Claim 38; Page 87; 116pp; English.
XX
CC AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC binding peptides are useful in modulating signal transduction pathways at
CC the cellular level (especially protein tyrosine kinase-mediated),
CC modulating oncogenic protein activity, or providing compounds for the
CC development of drugs with the ability to modulate broad classes, as well
CC as specific classes, of proteins involved in signal transduction and also
CC for regulating the processing, trafficking or translation of RNA.
CC Conjugates of the peptides with detectable labels or imaging agents are
CC useful for imaging cells, tissues and organs in which Src or Src-related
CC proteins are expressed
XX
SQ Sequence 13 AA;
Query Match 4.0%; Score 7; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 STRPLPP 112
Db 2 STRPLPP 8
RESULT 28
AAR93401
ID AAR93401 standard; peptide; 14 AA.
XX
AC AAR93401;
XX
DT 30-APR-1996 (first entry)
XX
DE GST-FYN SH3 protein tyrosine kinase derived peptide #11.
XX
KW SH3 ligand; SH3 binding agent; biased phage library;
KW recognition sequence; src SH3 domain; Paget's disease; restenosis;
KW rheumatoid arthritis; gout; neutrophil oxidase; tyrosine kinase; p47;
KW p67; complex; chronic myelogenous leukaemia; cancer.
```

```
XX
OS Synthetic.
XX
PN WO9524419-A1.
XX
PD 14-SEP-1995.
XX
PF 13-MAR-1995; 95WO-US003208.
XX
PR 11-MAR-1994; 94US-00209835.
PR 06-JAN-1995; 95US-00369832.
XX
PA (ARIA-) ARIAD PHARM INC.
XX
PI Rickles RJ, Brugge JS, Botfield MC, Zoller MJ;
XX
DR WPI; 1995-328231/42.
XX
PT Identification of peptide(s) binding specifically to SH3 domains - for
PT use in inhibiting interactions mediated by SH3 domains in treatment of
PT e.g. osteoporosis and cancer.
XX
PS Disclosure; Fig 4; 74pp; English.
XX
CC The sequences given in AAR93381-443 represent peptides which are SH3
CC ligands/SH3 binding agents. They represent a biased phage library which
CC comprises five random amino acids flanking the nonapeptide -RPLPLPPP
CC which was identified as a recognition sequence for the src SH3 domain.
CC These sequences were identified using the method of the invention. The
CC method comprises contacting the SH3 domain with a mixture of peptides
CC under conditions permitting a ligand to bind to an SH3 domain to form a
CC complex. Any unbound peptides are removed and the complexed peptide
CC ligands are dissociated from the complexes. The selected peptides are
CC enriched by re-contacting them with the SH3 domain and then candidates
CC which bind to the SH3 domain are detected. The isolated SH3 binding
CC peptides may be used in the diagnosis, prevention and treatment of
CC conditions or diseases resulting from cellular processes mediated by an
CC SH3-based interaction. Such diseases include Paget's disease. Other
CC conditions treatable with these peptides include restenosis, rheumatoid
CC arthritis, gout and other problems in which an SH3 of neutrophil oxidase
CC p47 and p67 complex is implicated, etc
XX
SQ Sequence 14 AA;
Query Match 4.0%; Score 7; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 STRPLPP 112
Db 4 STRPLPP 10
RESULT 29
AAW16947
ID AAW16947 standard; peptide; 31 AA.
XX
AC AAW16947;
XX
DT 27-JUN-1997 (first entry)
XX
DE Random recombinant SH3 domain binding peptide.
XX
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
```

PF 24-JUL-1995; 95WO-US009382.  
XX  
PR 22-JUL-1994; 94US-00278865.  
PR 07-JUN-1995; 95US-00483555.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA.  
XX  
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;  
XX WPI; 1996-117151/12.  
DR  
XX Peptide with binding affinity for Src homology region 3 (SH3) domains of  
PT proteins - useful for e.g. modulating signal transduction pathways at the  
PT cellular level, esp. protein tyrosine kinase-mediated.  
XX  
PS Disclosure; Fig 1; 116pp; English.  
XX  
CC AAW16924-W16948 are random recombinant peptides derived from one of three  
CC peptide libraries, T9, T12 and R8C. The peptides are all SH3 domain-  
CC binding peptides. SH3 binding peptides are useful in modulating signal  
CC transduction pathways at the cellular level (especially protein tyrosine  
CC kinase-mediated), modulating oncogenic protein activity, or providing  
CC compounds for the development of drugs with the ability to modulate broad  
CC classes, as well as specific classes, of proteins involved in signal  
CC transduction and also for regulating the processing, trafficking or  
CC translation of RNA. Conjugates of the peptides with detectable labels or  
CC imaging agents are useful for imaging cells, tissues and organs in which  
CC Src or Src-related proteins are expressed  
XX  
SQ Sequence 31 AA;  
  
Query Match 4.0%; Score 7; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 106 STRPLPP 112  
Db 16 STRPLPP 22  
  
RESULT 30  
AAW25510  
ID AAW25510 standard; peptide; 31 AA.  
XX  
AC AAW25510;  
XX  
DT 27-MAR-1998 (first entry)  
XX  
DE Random peptide recombinant clone R8C.YES3.8.  
XX  
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;  
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;  
KW PLCgamma; p53bp2; Crk; Yes; Grb2.  
XX  
OS Synthetic.  
OS Unidentified.  
XX  
PN WO9730074-A1.  
XX  
PD 21-AUG-1997.  
XX  
PF 14-FEB-1997; 97WO-US002298.  
XX  
PR 16-FEB-1996; 96US-00602999.  
XX  
XX (CYTO-) CYTOGEN CORP.  
PA (UYNC-) UNIV NORTH CAROLINA.  
XX  
XX Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;  
PI Rider JE;  
XX  
DR WPI; 1997-424972/39.  
XX

PT Src homology region 3 binding peptide - used to activate Src tyrosine  
PT kinase(s) and to stimulate immune response by increasing production of  
PT certain lymphokine(s), e.g. interleukin-1.  
XX  
PS Disclosure; Fig 5; 131pp; English.  
XX  
CC The present sequence represents a random peptide recombinant isolated by  
CC the method of the present invention. SH3 (src homology region 3) binding  
CC peptides are selected from: (a) peptides which bind the SH3 domain of  
CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)  
CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the  
CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;  
CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind  
CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3  
CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain  
CC of Grb2. The purified binding peptides can be used in the method to  
CC identify inhibitors of their binding to their respective SH3 domains,  
CC which could be used to modulate the pharmacological activity of proteins  
CC or polypeptide containing the SH3 domain. The peptides can also be used  
CC to activate Src or Src-related protein tyrosine kinases, to stimulate the  
CC immune response by increasing the production of certain lymphokines, e.g.  
CC tumour necrosis factor-alpha and interleukin-1, or to deliver a  
CC conjugated molecule to certain cellular compartments containing Src or  
CC Src related proteins  
XX  
SQ Sequence 31 AA;  
  
Query Match 4.0%; Score 7; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 106 STRPLPP 112  
Db 16 STRPLPP 22  
  
RESULT 31  
AAM20111  
ID AAM20111 standard; protein; 44 AA.  
XX  
AC AAM20111;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Peptide #6545 encoded by probe for measuring cervical gene expression.  
XX  
KW Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000670.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human cervical epithelial cells.



XX PS Claim 27; SEQ ID NO 24937; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes

CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded

CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

CC can be used to produce a single exon microarray, which can be used for

CC measuring human gene expression in a sample derived from human cervical

CC epithelial cells. By measuring gene expression, the probes are therefore

CC useful in grading and/or staging of diseases of the cervix, notably

CC cervical cancer. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 44 AA;

Query Match 4.0%; Score 7; DB 4; Length 44;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ALLGHRQ 64

Db |||||

37 ALLGHRQ 43

RESULT 32

ABB40309

ID ABB40309 standard; peptide; 44 AA.

XX AC ABB40309;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #7815 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human fetal liver.

PS Claim 27; SEQ ID NO 32944; 639pp + Sequence Listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and

CC displaying gene expression in samples derived from human fetal liver. The

CC present sequence is a peptide encoded by a single exon nucleic acid probe

CC of the invention. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 44 AA;

Query Match 4.0%; Score 7; DB 4; Length 44;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ALLGHRQ 64

Db |||||

37 ALLGHRQ 43

RESULT 33

AAM34001

ID AAM34001 standard; protein; 44 AA.

XX AC AAM34001;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #8038 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488897/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human placenta.

XX PS Claim 27; SEQ ID NO 34270; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENPs:

CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for producing a microarray for

CC predicting, measuring and displaying gene expression in samples derived

CC from human placenta. The probes are useful for antenatal diagnosis of

CC human genetic disorders

XX SQ Sequence 44 AA;

Query Match 4.0%; Score 7; DB 4; Length 44;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ALLGHRQ 64

Db |||||

37 ALLGHRQ 43

RESULT 34

ABB24712

ID ABB24712 standard; protein; 44 AA.

XX

AC ABB24712;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #6711 encoded by probe for measuring heart cell gene expression.  
XX  
KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
XX congenital heart disease.  
OS Homo sapiens.  
XX  
PN WO200157274-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000666.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488999/53.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts.  
XX  
PS Claim 15; SEQ ID NO 26482; 530pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC BY measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 44 AA;  
  
Query Match 4.0%; Score 7; DB 4; Length 44;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 58 ALLGHRQ 64  
Db 37 ALLGHRQ 43  
  
RESULT 35  
AAM73815  
ID AAM73815 standard; protein; 44 AA.  
XX  
AC AAM73815;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34121.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000668.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human bone marrow.  
XX  
PS Example 4; SEQ ID NO 34121; 658pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention  
XX  
SQ Sequence 44 AA;  
  
Query Match 4.0%; Score 7; DB 4; Length 44;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 58 ALLGHRQ 64  
Db 37 ALLGHRQ 43  
  
RESULT 36  
AAM61107  
ID AAM61107 standard; protein; 44 AA.  
XX  
AC AAM61107;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33212.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000667.  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.



PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483446/52.  
DR Single exon nucleic acid probes for analyzing gene expression in human  
XX brains.  
XX Example 4; SEQ ID NO 33212; 650pp + Sequence Listing; English.  
PS The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention  
XX  
SQ Sequence 44 AA;  
  
Query Match 4.0%; Score 7; DB 4; Length 44;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 58 ALLGHRQ 64  
Db 37 ALLGHRQ 43  
  
RESULT 37  
ABG55565  
ID ABG55565 standard; peptide; 44 AA.  
XX  
AC ABG55565;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID NO 34213.  
XX  
DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
KW  
XX Homo sapiens.  
OS  
XX  
PN WO200157273-A2.  
XX  
XX 09-AUG-2001.  
PD  
XX  
PF 30-JAN-2001; 2001WO-US0000664.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-488898/53.  
DR  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
PT  
XX Claim 27; SEQ ID NO 34213; 658pp; English.  
PS  
XX The invention relates to a single exon nucleic acid probe (SEN) (I) for

CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (I) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 44 AA;  
  
Query Match 4.0%; Score 7; DB 4; Length 44;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 58 ALLGHRQ 64  
Db 37 ALLGHRQ 43  
  
RESULT 38  
ABG43704  
ID ABG43704 standard; peptide; 44 AA.  
XX  
AC ABG43704;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 33369.  
DE  
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberos scleriosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
KW  
XX Homo sapiens.  
OS  
XX  
PN WO200186003-A2.  
XX  
XX 15-NOV-2001.  
PD  
XX  
PF 30-JAN-2001; 2001WO-US0000665.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2002-114183/15.  
DR  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples.  
PT  
XX Claim 27; SEQ ID NO 33369; 634pp; English.  
PS  
XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC ; the novel set of probes which hybridise at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC derived from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a peptide/protein encoded by a single exon probe of  
CC the invention. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 44 AA;

Query Match 4.0%; Score 7; DB 5; Length 44;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ALLGHRQ 64  
|||  
Db 37 ALLGHRQ 43

RESULT 39  
ABB17621

ID ABB17621 standard; protein; 50 AA.

XX AC ABB17621;

DT 23-JAN-2002 (first entry)

DE Human nervous system related polypeptide SEQ ID NO 5278.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antistickling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX OS Homo sapiens.

XX PN WO200159063-A2.

XX PD 16-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001334.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.



PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 20-OCT-2000; 2000US-0242221P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 01-DEC-2000; 2000US-0251160P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-541565/60.  
XX N-PSDB; ABA13947.

PT Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases.  
PT  
XX  
PS Claim 11; SEQ ID NO 6278; 1701pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins (ABB14678-ABB18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 50 AA;  
  
Query Match 4.0%; Score 7; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 46 CAVRAGL 52  
|||||||  
Db 21 CAVRAGL 27  
  
RESULT 40  
ABP07260  
ID ABP07260 standard; protein; 57 AA.  
XX  
AC ABP07260;  
XX  
DT 24-JUN-2002 (first entry)  
XX  
DE Human ORFX protein sequence SEQ ID NO:14502.  
XX  
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; degenerative disorder; osteoarthritis; neurodegenerative disorder; cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; hypertension; hypothyroidism; diabetes mellitus; cholesterl ester storage disease; immune deficiency; immune disorder; infectious disease; autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; myasthenia gravis.  
XX Homo sapiens.  
OS  
XX WO200192523-A2.  
PN  
XX 06-DEC-2001.  
PD  
XX 29-MAY-2001; 2001WO-US010836.  
PF  
XX 30-MAY-2000; 2000US-0206132P.  
PR 29-AUG-2000; 2000US-0228716P.  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX Shimkets RA, Leach MD;  
PI  
XX WPI; 2002-106308/14.  
DR N-PSDB; ABA13947.  
XX  
PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.

XX  
PS Disclosure; SEQ ID NO 14502; 1037pp; English.  
XX

CC The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic  
CC transplantation, cardiovascular diseases, disorders related to organ  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 57 AA;

Query Match 4.0%; Score 7; DB 5; Length 57;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
| | | | |  
Db 28 VAIVGNY 34

Search completed: May 18, 2004, 16:20:35  
Job time : 100 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:18:50 ; Search time 22 Seconds  
(without alignments)  
408.314 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 174  
Sequence: 1 RGHWVGATLGHNSRGFGV.....SAYASAQPTQPACPFPS 174

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued\_Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

*[Handwritten signature]*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	4.6	156	US-09-489-039A-11508	Sequence 11508, A
2	8	4.6	181	US-09-134-000C-4266	Sequence 4266, Ap
3	8	4.6	526	US-09-071-035-84	Sequence 84, Appl
4	8	4.6	546	US-09-071-035-82	Sequence 82, Appl
5	7	4.0	13	US-08-602-999A-93	Sequence 93, Appl
6	7	4.0	13	US-08-278-865-93	Sequence 93, Appl
7	7	4.0	13	US-09-500-124-93	Sequence 93, Appl
8	7	4.0	31	US-08-602-999A-60	Sequence 60, Appl
9	7	4.0	31	US-08-278-865-60	Sequence 60, Appl
10	7	4.0	31	US-09-500-124-60	Sequence 60, Appl
11	7	4.0	139	US-09-252-991A-23661	Sequence 23661, A
12	7	4.0	171	US-09-252-991A-26927	Sequence 26927, A
13	7	4.0	177	US-09-252-991A-18104	Sequence 18104, A
14	7	4.0	192	US-09-252-991A-21245	Sequence 21245, A
15	7	4.0	195	US-09-252-991A-27353	Sequence 27353, A
16	7	4.0	236	US-09-252-991A-27371	Sequence 27371, A
17	7	4.0	264	US-08-494-907-6	Sequence 6, Appli
18	7	4.0	264	PCT-US96-10986-6	Sequence 6018, A
19	7	4.0	363	US-09-252-991A-26018	Sequence 393, App
20	7	4.0	385	US-09-711-164-393	Sequence 22346, A
21	7	4.0	444	US-09-252-991A-22346	Sequence 22084, A
22	7	4.0	448	US-09-252-991A-22084	Sequence 28286, A
23	7	4.0	469	US-09-252-991A-28286	Sequence 9, Appli
24	7	4.0	544	US-08-591-685-9	Sequence 27527, A
25	7	4.0	760	US-09-252-991A-27527	Sequence 28773, A
26	7	4.0	1007	US-09-252-991A-28773	Sequence 82, Appl
27	6	3.4	13	US-08-602-999A-82	

28	6	3.4	13	3	US-08-602-999A-85	Sequence 85, Appl
29	6	3.4	13	3	US-08-602-999A-88	Sequence 88, Appl
30	6	3.4	13	4	US-08-278-865-82	Sequence 82, Appl
31	6	3.4	13	4	US-08-278-865-85	Sequence 85, Appl
32	6	3.4	13	4	US-08-278-865-88	Sequence 88, Appl
33	6	3.4	13	4	US-08-278-865-90	Sequence 90, Appl
34	6	3.4	13	4	US-09-500-124-82	Sequence 82, Appl
35	6	3.4	13	4	US-09-500-124-85	Sequence 85, Appl
36	6	3.4	13	4	US-09-500-124-88	Sequence 88, Appl
37	6	3.4	14	4	US-09-205-258-365	Sequence 365, App
38	6	3.4	15	3	US-08-602-999A-390	Sequence 390, App
39	6	3.4	15	4	US-09-500-124-390	Sequence 390, App
40	6	3.4	16	3	US-08-602-999A-411	Sequence 411, App
41	6	3.4	16	4	US-09-500-124-411	Sequence 411, App
42	6	3.4	17	3	US-08-602-999A-370	Sequence 370, App
43	6	3.4	17	4	US-09-500-124-370	Sequence 370, App
44	6	3.4	31	3	US-08-602-999A-53	Sequence 53, Appl
45	6	3.4	31	3	US-08-602-999A-56	Sequence 56, Appl
46	6	3.4	31	4	US-08-278-865-53	Sequence 53, Appl
47	6	3.4	31	4	US-08-278-865-56	Sequence 56, Appl
48	6	3.4	31	4	US-09-500-124-53	Sequence 53, Appl
49	6	3.4	31	4	US-09-500-124-56	Sequence 56, Appl
50	6	3.4	50	3	US-08-602-999A-50	Sequence 50, Appl
51	6	3.4	50	4	US-08-278-865-50	Sequence 50, Appl
52	6	3.4	50	4	US-09-500-124-50	Sequence 50, Appl
53	6	3.4	60	4	US-09-489-039A-10448	Sequence 10448, A
54	6	3.4	61	4	US-08-469-260A-468	Sequence 468, App
55	6	3.4	61	4	US-08-488-446-468	Sequence 468, App
56	6	3.4	61	4	US-08-467-344A-468	Sequence 468, App
57	6	3.4	73	1	US-08-274-366-3	Sequence 3, Appli
58	6	3.4	73	2	US-08-572-951-38	Sequence 38, Appl
59	6	3.4	73	2	US-08-392-816-19	Sequence 19, Appl
60	6	3.4	73	3	US-08-735-491-6	Sequence 6, Appli
61	6	3.4	73	5	PCT-US95-07828-3	Sequence 3, Appli
62	6	3.4	80	4	US-09-107-532A-3811	Sequence 3811, Ap
63	6	3.4	81	4	US-09-540-236-2674	Sequence 2674, Ap
64	6	3.4	82	4	US-09-621-976-5140	Sequence 5140, Ap
65	6	3.4	83	4	US-09-145-828A-2	Sequence 2, Appli
66	6	3.4	83	4	US-09-252-991A-26795	Sequence 26795, A
67	6	3.4	83	4	US-09-903-456-8	Sequence 8, Appli
68	6	3.4	85	4	US-09-252-991A-32003	Sequence 32003, A
69	6	3.4	87	4	US-09-252-991A-27237	Sequence 27237, A
70	6	3.4	91	4	US-10-164-595-75	Sequence 75, Appl
71	6	3.4	94	4	US-09-621-976-4685	Sequence 4685, Ap
72	6	3.4	96	4	US-09-252-991A-29153	Sequence 29153, A
73	6	3.4	99	4	US-09-621-976-5739	Sequence 5739, Ap
74	6	3.4	101	4	US-09-252-991A-29120	Sequence 29120, A
75	6	3.4	101	4	US-09-489-039A-8406	Sequence 8406, Ap
76	6	3.4	104	4	US-09-621-976-4162	Sequence 4162, Ap
77	6	3.4	106	4	US-09-252-991A-23993	Sequence 23993, A
78	6	3.4	108	4	US-09-252-991A-31220	Sequence 31220, A
79	6	3.4	113	4	US-09-252-991A-30749	Sequence 30749, A
80	6	3.4	115	4	US-09-621-976-5454	Sequence 5454, Ap
81	6	3.4	119	4	US-09-252-991A-25229	Sequence 25229, A
82	6	3.4	120	4	US-09-149-476-426	Sequence 426, App
83	6	3.4	123	4	US-09-107-532A-4158	Sequence 4158, Ap
84	6	3.4	124	4	US-09-252-991A-20183	Sequence 20183, A
85	6	3.4	127	4	US-09-489-039A-8110	Sequence 8110, Ap
86	6	3.4	128	4	US-09-489-039A-10886	Sequence 10886, A
87	6	3.4	132	4	US-09-454-034-2	Sequence 2, Appli
88	6	3.4	133	1	US-07-917-722-3	Sequence 3, Appli
89	6	3.4	134	4	US-09-252-991A-18017	Sequence 18017, A
90	6	3.4	140	4	US-09-252-991A-29232	Sequence 29232, A
91	6	3.4	143	4	US-09-252-991A-17322	Sequence 17322, A
92	6	3.4	144	2	US-09-108-051-2	Sequence 2, Appli
93	6	3.4	144	3	US-09-440-833-2	Sequence 2, Appli
94	6	3.4	146	4	US-09-690-454-63	Sequence 63, Appl
95	6	3.4	146	4	US-09-690-454-80	Sequence 80, Appl
96	6	3.4	147	4	US-09-252-991A-32004	Sequence 32004, A
97	6	3.4	150	4	US-09-732-210-626	Sequence 626, App
98	6	3.4	151	4	US-09-252-991A-23065	Sequence 23065, A
99	6	3.4	153	4	US-09-252-991A-30178	Sequence 30178, A
100	6	3.4	154	4	US-09-252-991A-20803	Sequence 20803, A

101	6	3.4	154	4	US-09-252-991A-21178	Sequence 21178, A
102	6	3.4	155	4	US-09-252-991A-26241	Sequence 26241, A
103	6	3.4	160	4	US-09-252-991A-22711	Sequence 22711, A
104	6	3.4	160	4	US-09-489-039A-10758	Sequence 10758, A
105	6	3.4	162	4	US-09-252-991A-31660	Sequence 31660, A
106	6	3.4	163	4	US-09-562-737-77	Sequence 77, Appl
107	6	3.4	164	4	US-09-252-991A-29980	Sequence 29980, A
108	6	3.4	168	4	US-09-252-991A-27872	Sequence 27872, A
109	6	3.4	170	4	US-09-252-991A-29833	Sequence 29833, A
110	6	3.4	171	4	US-09-252-991A-19210	Sequence 19210, A
111	6	3.4	173	4	US-09-252-991A-17716	Sequence 17716, A
112	6	3.4	174	4	US-09-621-976-5738	Sequence 5738, Ap
113	6	3.4	174	4	US-09-621-976-5746	Sequence 5746, Ap
114	6	3.4	178	4	US-09-056-556-236	Sequence 236, App
115	6	3.4	178	4	US-09-072-596-231	Sequence 231, App
116	6	3.4	178	4	US-09-252-991A-17000	Sequence 17000, A
117	6	3.4	178	4	US-09-252-991A-29942	Sequence 29942, A
118	6	3.4	178	4	US-09-072-967-236	Sequence 236, App
119	6	3.4	179	4	US-09-252-991A-27196	Sequence 27196, A
120	6	3.4	180	4	US-09-252-991A-17244	Sequence 17244, A
121	6	3.4	181	4	US-09-252-991A-26538	Sequence 26538, A
122	6	3.4	181	4	US-09-489-039A-14232	Sequence 14232, A
123	6	3.4	182	4	US-09-252-991A-32837	Sequence 32837, A
124	6	3.4	182	4	US-09-328-352-7468	Sequence 7468, Ap
125	6	3.4	182	4	US-09-673-395A-331	Sequence 331, App
126	6	3.4	184	4	US-09-252-991A-28187	Sequence 28187, A
127	6	3.4	187	4	US-09-252-991A-18109	Sequence 18109, A
128	6	3.4	188	4	US-09-252-991A-24358	Sequence 24358, A
129	6	3.4	188	4	US-09-252-991A-26365	Sequence 26365, A
130	6	3.4	189	1	US-07-982-650D-2	Sequence 2, Appli
131	6	3.4	189	1	US-08-331-379-2	Sequence 2, Appli
132	6	3.4	190	4	US-09-252-991A-21938	Sequence 21938, A
133	6	3.4	191	4	US-09-252-991A-21437	Sequence 21437, A
134	6	3.4	191	4	US-09-252-991A-30129	Sequence 30129, A
135	6	3.4	196	1	US-07-982-650D-1	Sequence 1, Appli
136	6	3.4	196	1	US-07-982-650D-3	Sequence 3, Appli
137	6	3.4	196	1	US-07-982-650D-4	Sequence 4, Appli
138	6	3.4	196	1	US-07-982-650D-5	Sequence 5, Appli
139	6	3.4	196	1	US-08-331-379-1	Sequence 1, Appli
140	6	3.4	196	1	US-08-331-379-3	Sequence 3, Appli
141	6	3.4	196	1	US-08-331-379-4	Sequence 4, Appli
142	6	3.4	196	1	US-08-331-379-5	Sequence 5, Appli
143	6	3.4	196	4	US-09-752-165-95	Sequence 95, Appl
144	6	3.4	198	4	US-09-325-932A-151	Sequence 151, App
145	6	3.4	199	2	US-08-900-407-4	Sequence 4, Appli
146	6	3.4	201	4	US-09-252-991A-23253	Sequence 23253, A
147	6	3.4	203	4	US-09-328-352-7071	Sequence 7071, Ap
148	6	3.4	205	4	US-09-489-039A-11450	Sequence 11450, A
149	6	3.4	207	4	US-09-252-991A-20975	Sequence 20975, A
150	6	3.4	212	4	US-09-252-991A-24512	Sequence 24512, A
151	6	3.4	214	4	US-09-489-039A-10885	Sequence 10885, A
152	6	3.4	217	4	US-09-252-991A-21077	Sequence 21077, A
153	6	3.4	217	4	US-09-198-452A-656	Sequence 656, App
154	6	3.4	221	4	US-09-252-991A-30144	Sequence 30144, A
155	6	3.4	224	4	US-09-252-991A-18944	Sequence 18944, A
156	6	3.4	224	4	US-09-252-991A-19251	Sequence 19251, A
157	6	3.4	233	4	US-09-252-991A-18018	Sequence 18018, A
158	6	3.4	237	4	US-09-252-991A-28116	Sequence 28116, A
159	6	3.4	238	4	US-09-107-532A-4828	Sequence 4828, Ap
160	6	3.4	242	4	US-09-198-452A-182	Sequence 182, App
161	6	3.4	244	4	US-09-489-039A-7623	Sequence 7623, Ap
162	6	3.4	249	2	US-08-626-685A-11	Sequence 11, Appl
163	6	3.4	252	4	US-09-252-991A-25346	Sequence 25346, A
164	6	3.4	253	4	US-09-489-039A-11603	Sequence 11603, A
165	6	3.4	257	4	US-09-252-991A-31359	Sequence 31359, A
166	6	3.4	261	4	US-09-252-991A-27270	Sequence 27270, A
167	6	3.4	261	4	US-09-252-991A-30655	Sequence 30655, A
168	6	3.4	261	4	US-09-252-991A-31043	Sequence 31043, A
169	6	3.4	268	4	US-09-461-325-346	Sequence 346, App
170	6	3.4	268	4	US-10-012-542-346	Sequence 346, App
171	6	3.4	270	4	US-09-399-913-59	Sequence 59, Appl
172	6	3.4	278	4	US-09-252-991A-28500	Sequence 28500, A
173	6	3.4	280	4	US-09-053-374A-5	Sequence 5, Appli

174	6	3.4	280	4	US-09-252-991A-27920	Sequence 27920, A
175	6	3.4	281	4	US-09-053-374A-2	Sequence 2, Appli
176	6	3.4	284	2	US-09-055-095-1	Sequence 1, Appli
177	6	3.4	285	4	US-09-023-942A-26	Sequence 26, Appl
178	6	3.4	286	4	US-09-252-991A-30343	Sequence 30343, A
179	6	3.4	288	4	US-09-252-991A-18436	Sequence 18436, A
180	6	3.4	288	4	US-09-252-991A-26185	Sequence 26185, A
181	6	3.4	290	4	US-09-252-991A-31977	Sequence 31977, A
182	6	3.4	292	4	US-09-252-991A-28308	Sequence 28308, A
183	6	3.4	293	4	US-09-252-991A-28184	Sequence 28184, A
184	6	3.4	294	4	US-09-252-991A-29716	Sequence 29716, A
185	6	3.4	295	4	US-09-489-039A-13961	Sequence 13961, A
186	6	3.4	302	4	US-09-252-991A-19798	Sequence 19798, A
187	6	3.4	305	4	US-09-252-991A-17593	Sequence 17593, A
188	6	3.4	310	3	US-08-796-792-4	Sequence 4, Appli
189	6	3.4	310	4	US-09-491-795-4	Sequence 4, Appli
190	6	3.4	310	4	US-09-907-794A-153	Sequence 153, App
191	6	3.4	310	4	US-09-905-125A-153	Sequence 153, App
192	6	3.4	310	4	US-09-902-775A-153	Sequence 153, App
193	6	3.4	311	4	US-09-252-991A-22295	Sequence 22295, A
194	6	3.4	311	4	US-09-252-991A-28676	Sequence 28676, A
195	6	3.4	311	4	US-09-489-039A-9759	Sequence 9759, Ap
196	6	3.4	312	3	US-09-230-637-34	Sequence 34, Appl
197	6	3.4	312	4	US-09-477-135A-134	Sequence 134, App
198	6	3.4	313	4	US-09-347-878-30	Sequence 30, Appl
199	6	3.4	313	4	US-09-367-007C-39	Sequence 39, Appl
200	6	3.4	314	3	US-09-188-930-193	Sequence 193, App
201	6	3.4	314	4	US-09-636-382A-2	Sequence 2, Appli
202	6	3.4	314	4	US-09-312-283C-193	Sequence 193, App
203	6	3.4	315	4	US-09-252-991A-28801	Sequence 28801, A
204	6	3.4	316	3	US-09-188-930-337	Sequence 337, App
205	6	3.4	316	4	US-09-312-283C-337	Sequence 337, App
206	6	3.4	318	4	US-09-489-039A-13319	Sequence 13319, A
207	6	3.4	323	4	US-09-252-991A-19527	Sequence 19527, A
208	6	3.4	323	4	US-09-252-991A-33113	Sequence 33113, A
209	6	3.4	324	4	US-09-252-991A-24664	Sequence 24664, A
210	6	3.4	326	1	US-08-571-758-11	Sequence 11, Appl
211	6	3.4	326	1	US-08-909-984A-11	Sequence 11, Appl
212	6	3.4	326	1	US-08-909-983-11	Sequence 11, Appl
213	6	3.4	326	4	US-09-345-473E-10	Sequence 10, Appl
214	6	3.4	326	4	US-09-134-000C-4813	Sequence 4813, Ap
215	6	3.4	329	4	US-09-252-991A-27332	Sequence 27332, A
216	6	3.4	330	3	US-09-145-391-2	Sequence 2, Appli
217	6	3.4	330	4	US-09-794-591-2	Sequence 2, Appli
218	6	3.4	331	2	US-08-997-080-182	Sequence 182, App
219	6	3.4	331	2	US-08-997-362-182	Sequence 182, App
220	6	3.4	331	3	US-09-095-855-182	Sequence 182, App
221	6	3.4	331	4	US-09-324-542-182	Sequence 182, App
222	6	3.4	331	4	US-09-205-426-182	Sequence 182, App
223	6	3.4	331	4	US-09-252-991A-28389	Sequence 28389, A
224	6	3.4	332	4	US-09-252-991A-30770	Sequence 30770, A
225	6	3.4	332	4	US-09-489-039A-8281	Sequence 8281, Ap
226	6	3.4	334	4	US-09-252-991A-32917	Sequence 32917, A
227	6	3.4	335	3	US-09-095-117-2	Sequence 2, Appli
228	6	3.4	335	3	US-09-095-117-4	Sequence 4, Appli
229	6	3.4	335	4	US-09-850-260-2	Sequence 2, Appli
230	6	3.4	335	4	US-09-850-260-4	Sequence 4, Appli
231	6	3.4	335	4	US-09-134-000C-4450	Sequence 4450, Ap
232	6	3.4	336	1	US-08-414-926A-26	Sequence 26, Appl
233	6	3.4	336	2	US-08-926-922-26	Sequence 26, Appl
234	6	3.4	336	3	US-09-253-682-26	Sequence 26, Appl
235	6	3.4	336	3	US-09-527-657-26	Sequence 26, Appl
236	6	3.4	336	4	US-09-543-681A-5611	Sequence 5611, Ap
237	6	3.4	336	4	US-09-892-100-26	Sequence 26, Appl
238	6	3.4	348	4	US-09-252-991A-25327	Sequence 25327, A
239	6	3.4	350	4	US-09-149-476-452	Sequence 452, App
240	6	3.4	353	4	US-09-489-039A-12270	Sequence 12270, A
241	6	3.4	353	4	US-09-134-000C-5843	Sequence 5843, Ap
242	6	3.4	356	3	US-09-232-191-4	Sequence 4, Appli
243	6	3.4	356	3	US-09-232-200-4	Sequence 4, Appli
244	6	3.4	356	4	US-09-232-197-4	Sequence 4, Appli
245	6	3.4	356	4	US-09-232-201-4	Sequence 4, Appli
246	6	3.4	356	4	US-09-232-195-4	Sequence 4, Appli



247	6	3.4	359	2	US-08-560-398-10	Sequence 10, Appl	320	6	3.4	450	4	US-09-252-991A-16659	Sequence 16659, A
248	6	3.4	360	4	US-09-252-991A-23756	Sequence 23756, A	321	6	3.4	458	4	US-09-252-991A-32991	Sequence 32991, A
249	6	3.4	360	4	US-09-252-991A-32452	Sequence 32452, A	322	6	3.4	459	4	US-09-513-783A-170	Sequence 170, App
250	6	3.4	360	4	US-09-543-681A-4783	Sequence 4783, Ap	323	6	3.4	459	4	US-09-252-991A-19393	Sequence 19393, A
251	6	3.4	360	4	US-09-489-039A-9383	Sequence 9383, Ap	324	6	3.4	460	4	US-09-489-039A-8966	Sequence 8966, Ap
252	6	3.4	362	3	US-09-120-365-74	Sequence 74, Appl	325	6	3.4	461	4	US-09-252-991A-27033	Sequence 27033, A
253	6	3.4	362	3	US-09-515-039-74	Sequence 74, Appl	326	6	3.4	463	4	US-09-252-991A-19381	Sequence 19381, A
254	6	3.4	362	4	US-09-252-991A-16985	Sequence 16985, A	327	6	3.4	465	4	US-09-252-991A-29633	Sequence 29633, A
255	6	3.4	365	4	US-09-252-991A-31439	Sequence 31439, A	328	6	3.4	475	4	US-09-252-991A-23581	Sequence 23581, A
256	6	3.4	366	4	US-09-252-991A-29361	Sequence 29361, A	329	6	3.4	476	4	US-09-252-991A-21580	Sequence 21580, A
257	6	3.4	371	4	US-09-252-991A-32719	Sequence 32719, A	330	6	3.4	477	4	US-09-252-991A-24617	Sequence 24617, A
258	6	3.4	372	2	US-08-626-685A-8	Sequence 8, Appli	331	6	3.4	481	4	US-09-252-991A-21926	Sequence 21926, A
259	6	3.4	372	3	US-08-993-088A-2	Sequence 2, Appli	332	6	3.4	481	4	US-09-252-991A-31662	Sequence 31662, A
260	6	3.4	372	3	US-08-993-088A-20	Sequence 20, Appl	333	6	3.4	487	4	US-09-252-991A-29392	Sequence 29392, A
261	6	3.4	372	4	US-08-993-424B-2	Sequence 2, Appli	334	6	3.4	489	2	US-08-752-307B-3	Sequence 3, Appli
262	6	3.4	372	4	US-08-665-034A-4	Sequence 2, Appli	335	6	3.4	489	3	US-09-330-317B-18	Sequence 18, Appli
263	6	3.4	372	4	US-08-665-034A-4	Sequence 4, Appli	336	6	3.4	489	4	US-09-707-802-3	Sequence 3, Appli
264	6	3.4	372	4	US-09-595-549-9	Sequence 9, Appli	337	6	3.4	489	4	US-09-991-326-3	Sequence 3, Appli
265	6	3.4	372	4	US-09-603-680-2	Sequence 2, Appli	338	6	3.4	489	4	US-09-808-589A-18	Sequence 18, Appl
266	6	3.4	372	4	US-09-603-680-20	Sequence 20, Appl	339	6	3.4	496	4	US-09-252-991A-21949	Sequence 21949, A
267	6	3.4	372	4	US-08-981-700A-2	Sequence 2, Appli	340	6	3.4	501	4	US-09-252-991A-25784	Sequence 25784, A
268	6	3.4	372	4	US-08-899-112B-8	Sequence 8, Appli	341	6	3.4	501	4	US-09-252-991A-31885	Sequence 31885, A
269	6	3.4	373	4	US-09-252-991A-29008	Sequence 29008, A	342	6	3.4	503	4	US-09-252-991A-23522	Sequence 23522, A
270	6	3.4	374	4	US-09-252-991A-21534	Sequence 21534, A	343	6	3.4	503	4	US-09-252-991A-32663	Sequence 32663, A
271	6	3.4	374	4	US-09-252-991A-30572	Sequence 30572, A	344	6	3.4	504	4	US-09-252-991A-28242	Sequence 28242, A
272	6	3.4	375	1	US-08-027-986-1	Sequence 1, Appli	345	6	3.4	505	4	US-09-252-991A-18964	Sequence 18964, A
273	6	3.4	375	1	US-08-027-986-2	Sequence 2, Appli	346	6	3.4	506	3	US-08-867-352-21	Sequence 21, Appl
274	6	3.4	376	4	US-09-252-991A-25057	Sequence 25057, A	347	6	3.4	506	4	US-09-360-237-59	Sequence 59, Appl
275	6	3.4	377	4	US-09-269-137-4	Sequence 4, Appli	348	6	3.4	510	4	US-09-252-991A-25076	Sequence 25076, A
276	6	3.4	380	4	US-09-461-325-161	Sequence 161, App	349	6	3.4	511	4	US-09-266-965-127	Sequence 127, App
277	6	3.4	380	4	US-10-012-542-161	Sequence 161, App	350	6	3.4	512	4	US-09-252-991A-30763	Sequence 30763, A
278	6	3.4	385	1	US-08-539-798-4	Sequence 4, Appli	351	6	3.4	514	4	US-09-252-991A-22462	Sequence 22462, A
279	6	3.4	385	1	US-08-329-560-4	Sequence 4, Appli	352	6	3.4	518	4	US-09-252-991A-23604	Sequence 23604, A
280	6	3.4	388	4	US-09-266-965-10	Sequence 10, Appl	353	6	3.4	523	4	US-09-252-991A-19841	Sequence 19841, A
281	6	3.4	390	4	US-09-489-039A-13547	Sequence 13547, A	354	6	3.4	527	4	US-09-489-039A-10883	Sequence 10883, A
282	6	3.4	399	4	US-09-134-000C-3786	Sequence 3786, Ap	355	6	3.4	528	1	US-08-368-071-13	Sequence 13, Appl
283	6	3.4	401	4	US-09-252-991A-22213	Sequence 22213, A	356	6	3.4	528	1	US-08-458-181-13	Sequence 13, Appl
284	6	3.4	403	2	US-08-812-412-2	Sequence 2, Appli	357	6	3.4	530	5	PCT-US93-02172-13	Sequence 2, Appli
285	6	3.4	403	3	US-09-180-271-5	Sequence 5, Appli	358	6	3.4	530	2	US-08-752-307B-2	Sequence 2, Appli
286	6	3.4	405	3	US-09-232-200-63	Sequence 63, Appl	359	6	3.4	530	4	US-09-707-802-2	Sequence 2, Appli
287	6	3.4	405	4	US-09-232-197-63	Sequence 63, Appl	360	6	3.4	530	4	US-09-991-326-2	Sequence 2, Appli
288	6	3.4	405	4	US-09-232-201-63	Sequence 63, Appl	361	6	3.4	531	4	US-09-252-991A-27660	Sequence 27660, A
289	6	3.4	405	4	US-09-252-991A-25946	Sequence 25946, A	362	6	3.4	532	4	US-09-252-991A-19241	Sequence 19241, A
290	6	3.4	405	4	US-09-232-195-63	Sequence 63, Appl	363	6	3.4	534	4	US-09-252-991A-17265	Sequence 17265, A
291	6	3.4	406	4	US-09-252-991A-19857	Sequence 19857, A	364	6	3.4	538	4	US-09-252-991A-32064	Sequence 32064, A
292	6	3.4	407	4	US-09-252-991A-32423	Sequence 32423, A	365	6	3.4	548	4	US-09-107-532A-6627	Sequence 6627, Ap
293	6	3.4	407	4	US-09-252-991A-18004	Sequence 18004, A	366	6	3.4	550	4	US-09-252-991A-25062	Sequence 25062, A
294	6	3.4	410	4	US-09-252-991A-26318	Sequence 26318, A	367	6	3.4	550	4	US-09-252-991A-27117	Sequence 27117, A
295	6	3.4	410	4	US-09-252-991A-31487	Sequence 31487, A	368	6	3.4	553	4	US-09-252-991A-17089	Sequence 17089, A
296	6	3.4	411	4	US-09-252-991A-22445	Sequence 22445, A	369	6	3.4	554	4	US-09-489-039A-10541	Sequence 10541, A
297	6	3.4	411	4	US-09-252-991A-31301	Sequence 31301, A	370	6	3.4	558	4	US-09-351-150A-7	Sequence 7, Appli
298	6	3.4	413	4	US-09-489-039A-7856	Sequence 7856, Ap	371	6	3.4	560	2	US-08-928-692-53	Sequence 53, Appl
299	6	3.4	414	4	US-09-252-991A-26863	Sequence 26863, A	372	6	3.4	560	4	US-09-339-972-53	Sequence 53, Appl
300	6	3.4	415	4	US-09-252-991A-31684	Sequence 31684, A	373	6	3.4	565	4	US-09-252-991A-32570	Sequence 32570, A
301	6	3.4	416	2	US-08-882-704A-7	Sequence 7, Appli	374	6	3.4	578	4	US-09-107-532A-5883	Sequence 5883, Ap
302	6	3.4	416	4	US-09-151-957-7	Sequence 7, Appli	375	6	3.4	579	4	US-09-252-991A-24120	Sequence 24120, A
303	6	3.4	416	4	US-09-328-352-6537	Sequence 6537, Ap	376	6	3.4	581	4	US-10-023-515-2	Sequence 2, Appli
304	6	3.4	416	4	US-10-195-158-7	Sequence 7, Appli	377	6	3.4	583	3	US-08-941-445A-13	Sequence 13, Appl
305	6	3.4	417	3	US-08-640-906-4	Sequence 4, Appli	378	6	3.4	586	4	US-09-134-001C-4456	Sequence 4456, Ap
306	6	3.4	417	3	US-08-640-906-18	Sequence 18, Appl	379	6	3.4	597	4	US-09-252-991A-31342	Sequence 31342, A
307	6	3.4	417	4	US-09-395-936-4	Sequence 4, Appli	380	6	3.4	599	4	US-09-252-991A-20985	Sequence 20985, A
308	6	3.4	417	4	US-09-395-936-18	Sequence 18, Appl	381	6	3.4	611	4	US-09-543-681A-7786	Sequence 7786, Ap
309	6	3.4	418	4	US-09-252-991A-31753	Sequence 31753, A	382	6	3.4	614	4	US-09-252-991A-29695	Sequence 29695, A
310	6	3.4	430	1	US-08-601-435-2	Sequence 2, Appli	383	6	3.4	616	4	US-09-252-991A-25638	Sequence 25638, A
311	6	3.4	430	2	US-08-931-047-2	Sequence 2, Appli	384	6	3.4	632	3	US-09-232-200-34	Sequence 34, Appl
312	6	3.4	430	2	US-08-783-202-2	Sequence 2, Appli	385	6	3.4	632	3	US-09-232-200-39	Sequence 39, Appl
313	6	3.4	430	4	US-09-443-041A-31	Sequence 31, Appl	386	6	3.4	632	4	US-09-232-197-34	Sequence 34, Appl
314	6	3.4	432	4	US-09-443-041A-20	Sequence 20, Appl	387	6	3.4	632	4	US-09-232-197-39	Sequence 39, Appl
315	6	3.4	440	4	US-09-252-991A-24173	Sequence 24173, A	388	6	3.4	632	4	US-09-232-201-34	Sequence 34, Appl
316	6	3.4	440	4	US-09-252-991A-24174	Sequence 24174, A	389	6	3.4	632	4	US-09-232-201-39	Sequence 39, Appl
317	6	3.4	440	4	US-09-543-681A-7002	Sequence 7002, Ap	390	6	3.4	632	4	US-09-661-711A-5	Sequence 5, Appli
318	6	3.4	441	4	US-09-252-991A-32115	Sequence 32115, A	391	6	3.4	632	4	US-09-232-195-34	Sequence 34, Appl
319	6	3.4	443	4	US-09-489-039A-8166	Sequence 8166, Ap	392	6	3.4	632	4	US-09-232-195-39	Sequence 39, Appl

393	6	3.4	639	4	US-09-252-991A-18903	Sequence 18903, A	466	6	3.4	762	4	US-09-579-365-2	Sequence 2, Appli
394	6	3.4	639	4	US-09-252-991A-23693	Sequence 23693, A	467	6	3.4	790	4	US-09-252-991A-23247	Sequence 23247, A
395	6	3.4	639	4	US-09-252-991A-25089	Sequence 25089, A	468	6	3.4	806	4	US-08-684-005-2	Sequence 2, Appli
396	6	3.4	639	4	US-10-164-595-44	Sequence 44, Appl	469	6	3.4	817	4	US-09-252-991A-25887	Sequence 25887, A
397	6	3.4	643	3	US-09-232-200-27	Sequence 27, Appl	470	6	3.4	821	4	US-09-252-991A-30347	Sequence 30347, A
398	6	3.4	643	3	US-09-232-200-41	Sequence 41, Appl	471	6	3.4	885	4	US-09-252-991A-26129	Sequence 26129, A
399	6	3.4	643	3	US-09-232-200-53	Sequence 53, Appl	472	6	3.4	887	4	US-09-489-039A-11718	Sequence 11718, A
400	6	3.4	643	4	US-09-232-197-27	Sequence 27, Appl	473	6	3.4	896	4	US-09-543-681A-5439	Sequence 5439, Ap
401	6	3.4	643	4	US-09-232-197-41	Sequence 41, Appl	474	6	3.4	927	4	US-09-198-452A-472	Sequence 472, App
402	6	3.4	643	4	US-09-232-197-53	Sequence 53, Appl	475	6	3.4	966	4	US-09-252-991A-24356	Sequence 24356, A
403	6	3.4	643	4	US-09-232-201-27	Sequence 27, Appl	476	6	3.4	1036	4	US-09-489-039A-10266	Sequence 10266, A
404	6	3.4	643	4	US-09-232-201-41	Sequence 41, Appl	477	6	3.4	1073	4	US-09-252-991A-30317	Sequence 30317, A
405	6	3.4	643	4	US-09-232-201-53	Sequence 53, Appl	478	6	3.4	1124	4	US-08-311-731A-10	Sequence 10, Appl
406	6	3.4	643	4	US-09-232-195-27	Sequence 27, Appl	479	6	3.4	1133	4	US-09-252-991A-32131	Sequence 32131, A
407	6	3.4	643	4	US-09-232-195-41	Sequence 41, Appl	480	6	3.4	1138	4	US-09-252-991A-20291	Sequence 20291, A
408	6	3.4	643	4	US-09-232-195-53	Sequence 53, Appl	481	6	3.4	1140	3	US-09-220-081-2	Sequence 2, Appli
409	6	3.4	646	3	US-09-232-200-25	Sequence 25, Appl	482	6	3.4	1140	4	US-09-677-575-2	Sequence 2, Appli
410	6	3.4	646	3	US-09-232-200-32	Sequence 32, Appl	483	6	3.4	1296	3	US-08-728-603-15	Sequence 15, Appl
411	6	3.4	646	3	US-09-232-200-33	Sequence 33, Appl	484	6	3.4	1421	3	US-09-335-409-2	Sequence 2, Appli
412	6	3.4	646	3	US-09-232-200-38	Sequence 38, Appl	485	6	3.4	1421	4	US-09-568-102-2	Sequence 2, Appli
413	6	3.4	646	3	US-09-232-200-43	Sequence 43, Appl	486	6	3.4	1421	4	US-09-567-969-2	Sequence 2, Appli
414	6	3.4	646	3	US-09-232-200-47	Sequence 47, Appl	487	6	3.4	1421	4	US-09-568-480-2	Sequence 2, Appli
415	6	3.4	646	3	US-09-232-200-59	Sequence 59, Appl	488	6	3.4	1421	4	US-09-568-486-2	Sequence 2, Appli
416	6	3.4	646	3	US-09-232-200-65	Sequence 65, Appl	489	6	3.4	1421	4	US-09-568-472-2	Sequence 2, Appli
417	6	3.4	646	3	US-09-232-200-92	Sequence 92, Appl	490	6	3.4	1421	4	US-09-567-899-2	Sequence 2, Appli
418	6	3.4	646	4	US-09-232-197-25	Sequence 25, Appl	491	6	3.4	1452	4	US-09-127-227-2	Sequence 2, Appli
419	6	3.4	646	4	US-09-232-197-32	Sequence 32, Appl	492	6	3.4	1493	4	US-09-489-039A-13687	Sequence 13687, A
420	6	3.4	646	4	US-09-232-197-33	Sequence 33, Appl	493	6	3.4	1503	4	US-09-600-087-2	Sequence 2, Appli
421	6	3.4	646	4	US-09-232-197-38	Sequence 38, Appl	494	6	3.4	1593	4	US-08-628-829-4	Sequence 4, Appli
422	6	3.4	646	4	US-09-232-197-43	Sequence 43, Appl	495	6	3.4	1692	3	US-09-263-933-4	Sequence 4, Appli
423	6	3.4	646	4	US-09-232-197-47	Sequence 47, Appl	496	6	3.4	1692	3	US-09-263-933-11	Sequence 11, Appl
424	6	3.4	646	4	US-09-232-197-59	Sequence 59, Appl	497	6	3.4	1692	3	US-09-263-933-18	Sequence 18, Appl
425	6	3.4	646	4	US-09-232-197-65	Sequence 65, Appl	498	6	3.4	1692	4	US-09-919-901-4	Sequence 4, Appli
426	6	3.4	646	4	US-09-232-197-92	Sequence 92, Appl	499	6	3.4	1692	4	US-09-919-901-11	Sequence 11, Appl
427	6	3.4	646	4	US-09-232-201-25	Sequence 25, Appl	500	6	3.4	1692	4	US-09-919-901-18	Sequence 18, Appl
428	6	3.4	646	4	US-09-232-201-32	Sequence 32, Appl	501	6	3.4	1702	3	US-08-296-791-5	Sequence 5, Appli
429	6	3.4	646	4	US-09-232-201-33	Sequence 33, Appl	502	6	3.4	1702	4	US-09-839-996-5	Sequence 5, Appli
430	6	3.4	646	4	US-09-232-201-38	Sequence 38, Appl	503	6	3.4	1702	4	US-10-080-505-5	Sequence 5, Appli
431	6	3.4	646	4	US-09-232-201-43	Sequence 43, Appl	504	6	3.4	1702	5	PCT-US95-10661A-5	Sequence 5, Appli
432	6	3.4	646	4	US-09-232-201-47	Sequence 47, Appl	505	6	3.4	1706	4	US-09-252-991A-31760	Sequence 31760, A
433	6	3.4	646	4	US-09-232-201-59	Sequence 59, Appl	506	6	3.4	1848	3	US-08-296-791-6	Sequence 6, Appli
434	6	3.4	646	4	US-09-232-201-65	Sequence 65, Appl	507	6	3.4	1848	4	US-09-839-996-6	Sequence 6, Appli
435	6	3.4	646	4	US-09-232-201-92	Sequence 92, Appl	508	6	3.4	1848	4	US-10-080-505-6	Sequence 6, Appli
436	6	3.4	646	4	US-09-232-195-25	Sequence 25, Appl	509	6	3.4	1848	5	PCT-US95-10661A-6	Sequence 6, Appli
437	6	3.4	646	4	US-09-232-195-32	Sequence 32, Appl	510	6	3.4	1891	2	US-08-804-227C-12	Sequence 12, Appl
438	6	3.4	646	4	US-09-232-195-33	Sequence 33, Appl	511	6	3.4	1891	2	US-08-804-198-6	Sequence 6, Appli
439	6	3.4	646	4	US-09-232-195-38	Sequence 38, Appl	512	6	3.4	2307	3	US-09-263-933-2	Sequence 2, Appli
440	6	3.4	646	4	US-09-232-195-43	Sequence 43, Appl	513	6	3.4	2307	3	US-09-263-933-9	Sequence 9, Appli
441	6	3.4	646	4	US-09-232-195-47	Sequence 47, Appl	514	6	3.4	2307	3	US-09-263-933-16	Sequence 16, Appl
442	6	3.4	646	4	US-09-232-195-59	Sequence 59, Appl	515	6	3.4	2307	4	US-09-919-901-2	Sequence 2, Appli
443	6	3.4	646	4	US-09-232-195-65	Sequence 65, Appl	516	6	3.4	2307	4	US-09-919-901-9	Sequence 9, Appli
444	6	3.4	646	4	US-09-232-195-92	Sequence 92, Appl	517	6	3.4	2307	4	US-09-919-901-16	Sequence 16, Appl
445	6	3.4	651	4	US-09-252-991A-32204	Sequence 32204, A	518	6	3.4	2385	4	US-09-543-681A-6304	Sequence 6304, Ap
446	6	3.4	660	4	US-09-976-594-787	Sequence 787, App	519	6	3.4	2474	4	US-08-265-967C-3	Sequence 3, Appli
447	6	3.4	661	4	US-09-252-991A-29083	Sequence 29083, A	520	6	3.4	2474	4	US-08-305-790B-4	Sequence 4, Appli
448	6	3.4	672	4	US-09-252-991A-18930	Sequence 18930, A	521	6	3.4	2628	3	US-09-413-814-11	Sequence 11, Appl
449	6	3.4	672	4	US-09-252-991A-32554	Sequence 32554, A	522	6	3.4	3170	2	US-07-642-734C-5	Sequence 5, Appli
450	6	3.4	685	4	US-09-252-991A-32033	Sequence 32033, A	523	6	3.4	3170	3	US-08-439-009A-5	Sequence 5, Appli
451	6	3.4	690	4	US-09-252-991A-29429	Sequence 29429, A	524	6	3.4	3457	2	US-08-416-603-4	Sequence 4, Appli
452	6	3.4	697	4	US-09-252-991A-21106	Sequence 21106, A	525	6	3.4	3562	4	US-09-679-279-14	Sequence 14, Appl
453	6	3.4	705	4	US-09-252-991A-27442	Sequence 27442, A	526	6	3.4	3567	2	US-07-642-734C-4	Sequence 4, Appli
454	6	3.4	711	4	US-09-252-991A-28199	Sequence 28199, A	527	6	3.4	3567	3	US-08-439-009A-4	Sequence 4, Appli
455	6	3.4	713	4	US-09-907-794A-245	Sequence 245, App	528	6	3.4	3665	2	US-08-222-617A-13	Sequence 13, Appl
456	6	3.4	713	4	US-09-905-125A-245	Sequence 245, App	529	6	3.4	3712	2	US-08-222-617A-4	Sequence 4, Appli
457	6	3.4	713	4	US-09-902-775A-245	Sequence 245, App	530	6	3.4	3712	2	US-08-222-617A-25	Sequence 25, Appl
458	6	3.4	718	4	US-09-252-991A-31865	Sequence 31865, A	531	6	3.4	3729	2	US-08-804-227C-4	Sequence 4, Appli
459	6	3.4	719	4	US-09-252-991A-22278	Sequence 22278, A	532	6	3.4	4654	3	US-08-476-515A-84	Sequence 84, Appl
460	6	3.4	722	4	US-09-252-991A-22811	Sequence 22811, A	533	6	3.4	4655	3	US-08-652-877-84	Sequence 84, Appl
461	6	3.4	728	4	US-09-747-259-18	Sequence 18, Appl	534	6	3.4	4655	3	US-08-652-877-86	Sequence 86, Appl
462	6	3.4	728	4	US-09-816-744-18	Sequence 18, Appl	535	6	3.4	4655	3	US-08-652-877-88	Sequence 88, Appl
463	6	3.4	743	3	US-09-077-354B-2	Sequence 2, Appli	536	6	3.4	4655	3	US-08-652-877-90	Sequence 90, Appl
464	6	3.4	744	4	US-09-252-991A-19290	Sequence 19290, A	537	6	3.4	7257	3	US-09-335-409-5	Sequence 5, Appli
465	6	3.4	756	4	US-09-252-991A-30809	Sequence 30809, A	538	6	3.4	7257	4	US-09-568-102-5	Sequence 5, Appli



539	6	3.4	7257	4	US-09-567-969-5	Sequence 5, Appli	612	5	2.9	12	2	US-08-911-092-26	Sequence 26, Appl
540	6	3.4	7257	4	US-09-568-480-5	Sequence 5, Appli	613	5	2.9	12	2	US-08-485-001B-13	Sequence 13, Appl
541	6	3.4	7257	4	US-09-568-486-5	Sequence 5, Appli	614	5	2.9	12	2	US-08-485-001B-14	Sequence 14, Appl
542	6	3.4	7257	4	US-09-568-472-5	Sequence 5, Appli	615	5	2.9	12	2	US-08-485-001B-15	Sequence 15, Appl
543	6	3.4	7257	4	US-09-567-899-5	Sequence 5, Appli	616	5	2.9	12	2	US-08-485-001B-21	Sequence 21, Appl
544	5	2.9	7	1	US-08-487-890A-78	Sequence 78, Appl	617	5	2.9	12	2	US-08-485-001B-22	Sequence 22, Appl
545	5	2.9	7	2	US-08-478-435-78	Sequence 78, Appl	618	5	2.9	12	2	US-08-485-001B-23	Sequence 23, Appl
546	5	2.9	7	2	US-08-337-483-78	Sequence 78, Appl	619	5	2.9	12	2	US-08-485-001B-24	Sequence 24, Appl
547	5	2.9	7	2	US-08-478-373-78	Sequence 78, Appl	620	5	2.9	12	2	US-08-485-001B-25	Sequence 25, Appl
548	5	2.9	7	2	US-08-680-326-85	Sequence 85, Appl	621	5	2.9	12	2	US-08-485-001B-26	Sequence 26, Appl
549	5	2.9	7	2	US-08-769-745-7	Sequence 7, Appli	622	5	2.9	12	3	US-08-454-121A-13	Sequence 13, Appl
550	5	2.9	7	2	US-08-769-745-26	Sequence 26, Appl	623	5	2.9	12	3	US-08-454-121A-14	Sequence 14, Appl
551	5	2.9	7	3	US-08-474-671-78	Sequence 78, Appl	624	5	2.9	12	3	US-08-454-121A-15	Sequence 15, Appl
552	5	2.9	7	3	US-08-483-577A-78	Sequence 78, Appl	625	5	2.9	12	3	US-08-454-121A-21	Sequence 21, Appl
553	5	2.9	7	3	US-08-602-999A-9	Sequence 9, Appli	626	5	2.9	12	3	US-08-454-121A-22	Sequence 22, Appl
554	5	2.9	7	3	US-08-897-438-78	Sequence 78, Appl	627	5	2.9	12	3	US-08-454-121A-23	Sequence 23, Appl
555	5	2.9	7	4	US-08-278-865-9	Sequence 9, Appli	628	5	2.9	12	3	US-08-454-121A-24	Sequence 24, Appl
556	5	2.9	7	4	US-08-630-915A-45	Sequence 45, Appl	629	5	2.9	12	3	US-08-454-121A-25	Sequence 25, Appl
557	5	2.9	7	4	US-08-637-654-78	Sequence 78, Appl	630	5	2.9	12	3	US-08-454-121A-26	Sequence 26, Appl
558	5	2.9	7	4	US-08-649-518-78	Sequence 78, Appl	631	5	2.9	12	3	US-08-482-161B-13	Sequence 13, Appl
559	5	2.9	7	4	US-09-500-124-9	Sequence 9, Appli	632	5	2.9	12	3	US-08-482-161B-14	Sequence 14, Appl
560	5	2.9	7	4	US-09-428-082B-282	Sequence 282, App	633	5	2.9	12	3	US-08-482-161B-15	Sequence 15, Appl
561	5	2.9	7	4	US-09-428-082B-303	Sequence 303, App	634	5	2.9	12	3	US-08-482-161B-21	Sequence 21, Appl
562	5	2.9	7	4	US-09-428-082B-323	Sequence 323, App	635	5	2.9	12	3	US-08-482-161B-22	Sequence 22, Appl
563	5	2.9	8	4	US-08-952-445-11	Sequence 11, Appl	636	5	2.9	12	3	US-08-482-161B-23	Sequence 23, Appl
564	5	2.9	9	1	US-08-465-167A-34	Sequence 34, Appl	637	5	2.9	12	3	US-08-482-161B-24	Sequence 24, Appl
565	5	2.9	9	1	US-08-615-181-28	Sequence 28, Appl	638	5	2.9	12	3	US-08-482-161B-25	Sequence 25, Appl
566	5	2.9	9	2	US-08-126-016-13	Sequence 13, Appl	639	5	2.9	12	3	US-08-482-161B-26	Sequence 26, Appl
567	5	2.9	9	2	US-08-126-016-22	Sequence 22, Appl	640	5	2.9	12	3	US-09-057-963A-11	Sequence 11, Appl
568	5	2.9	9	3	US-08-602-999A-11	Sequence 11, Appl	641	5	2.9	12	3	US-09-057-963A-12	Sequence 12, Appl
569	5	2.9	9	4	US-08-278-865-11	Sequence 11, Appl	642	5	2.9	12	3	US-09-057-963A-13	Sequence 13, Appl
570	5	2.9	9	4	US-09-302-305C-14	Sequence 14, Appl	643	5	2.9	12	3	US-09-057-963A-21	Sequence 21, Appl
571	5	2.9	9	4	US-09-302-305C-27	Sequence 27, Appl	644	5	2.9	12	3	US-09-057-963A-22	Sequence 22, Appl
572	5	2.9	9	4	US-09-500-124-11	Sequence 11, Appl	645	5	2.9	12	3	US-09-057-963A-23	Sequence 23, Appl
573	5	2.9	10	1	US-08-346-333-62	Sequence 62, Appl	646	5	2.9	12	3	US-09-057-963A-24	Sequence 24, Appl
574	5	2.9	10	1	US-08-346-333-68	Sequence 68, Appl	647	5	2.9	12	3	US-09-057-963A-25	Sequence 25, Appl
575	5	2.9	10	2	US-08-556-597-87	Sequence 87, Appl	648	5	2.9	12	3	US-09-057-963A-26	Sequence 26, Appl
576	5	2.9	10	2	US-08-468-812-10	Sequence 10, Appl	649	5	2.9	12	3	US-09-057-963A-26	Sequence 307, App
577	5	2.9	10	2	US-08-769-745-5	Sequence 5, Appli	650	5	2.9	12	4	US-09-428-082B-307	Sequence 308, App
578	5	2.9	10	3	US-08-159-339A-520	Sequence 520, App	651	5	2.9	12	4	US-09-428-082B-308	Sequence 309, App
579	5	2.9	10	3	US-08-836-075A-193	Sequence 193, App	652	5	2.9	12	4	US-09-428-082B-309	Sequence 310, App
580	5	2.9	10	3	US-09-461-697-9	Sequence 9, Appli	653	5	2.9	12	4	US-09-428-082B-310	Sequence 311, App
581	5	2.9	10	4	US-08-590-563-18	Sequence 18, Appl	654	5	2.9	12	5	PCT-US94-10257A-19	Sequence 19, Appl
582	5	2.9	10	4	US-09-770-621-18	Sequence 18, Appl	655	5	2.9	12	6	5187077-33	Patent No. 5187077
583	5	2.9	10	4	US-09-311-784A-369	Sequence 369, App	656	5	2.9	12	6	5427925-31	Patent No. 5427925
584	5	2.9	10	4	US-09-235-832-18	Sequence 18, Appl	657	5	2.9	13	3	US-08-855-531D-40	Sequence 40, Appl
585	5	2.9	10	5	PCT-US91-07506-62	Sequence 62, Appl	658	5	2.9	13	3	US-08-974-549A-186	Sequence 186, App
586	5	2.9	10	5	PCT-US91-07506-68	Sequence 68, Appl	659	5	2.9	13	3	US-08-602-999A-66	Sequence 66, Appl
587	5	2.9	11	1	US-08-336-343A-24	Sequence 24, Appl	660	5	2.9	13	3	US-08-602-999A-81	Sequence 81, Appl
588	5	2.9	11	3	US-08-652-877-34	Sequence 34, Appl	661	5	2.9	13	3	US-08-855-526B-40	Sequence 40, Appl
589	5	2.9	11	3	US-08-652-877-35	Sequence 35, Appl	662	5	2.9	13	4	US-08-278-865-66	Sequence 66, Appl
590	5	2.9	11	3	US-08-652-877-36	Sequence 36, Appl	663	5	2.9	13	4	US-08-278-865-81	Sequence 81, Appl
591	5	2.9	11	3	US-08-476-515A-34	Sequence 34, Appl	664	5	2.9	13	4	US-08-630-915A-179	Sequence 179, App
592	5	2.9	11	3	US-08-476-515A-35	Sequence 35, Appl	665	5	2.9	13	4	US-08-630-915A-187	Sequence 187, App
593	5	2.9	11	3	US-08-476-515A-36	Sequence 36, Appl	666	5	2.9	13	4	US-08-630-915A-202	Sequence 202, App
594	5	2.9	11	4	US-09-025-769B-186	Sequence 186, App	667	5	2.9	13	4	US-09-500-124-66	Sequence 66, Appl
595	5	2.9	12	2	US-08-489-666C-13	Sequence 13, Appl	668	5	2.9	13	4	US-09-500-124-81	Sequence 81, Appl
596	5	2.9	12	2	US-08-489-666C-14	Sequence 14, Appl	669	5	2.9	13	4	US-08-469-260A-590	Sequence 590, App
597	5	2.9	12	2	US-08-489-666C-15	Sequence 15, Appl	670	5	2.9	13	4	US-08-488-446-590	Sequence 590, App
598	5	2.9	12	2	US-08-489-666C-21	Sequence 21, Appl	671	5	2.9	13	4	US-08-467-344A-590	Sequence 590, App
599	5	2.9	12	2	US-08-489-666C-22	Sequence 22, Appl	672	5	2.9	13	4	US-09-402-181B-186	Sequence 186, App
600	5	2.9	12	2	US-08-489-666C-23	Sequence 23, Appl	673	5	2.9	13	4	US-09-721-456-186	Sequence 186, App
601	5	2.9	12	2	US-08-489-666C-24	Sequence 24, Appl	674	5	2.9	13	4	US-09-528-603-8	Sequence 8, Appli
602	5	2.9	12	2	US-08-489-666C-25	Sequence 25, Appl	675	5	2.9	13	5	PCT-US95-04121-29	Sequence 29, Appl
603	5	2.9	12	2	US-08-489-666C-26	Sequence 26, Appl	676	5	2.9	13	5	PCT-US95-04121-30	Sequence 30, Appl
604	5	2.9	12	2	US-08-911-092-13	Sequence 13, Appl	677	5	2.9	14	2	US-08-480-190-51	Sequence 51, Appl
605	5	2.9	12	2	US-08-911-092-14	Sequence 14, Appl	678	5	2.9	14	2	US-08-488-379-51	Sequence 51, Appl
606	5	2.9	12	2	US-08-911-092-15	Sequence 15, Appl	679	5	2.9	14	3	US-08-946-329A-13	Sequence 13, Appl
607	5	2.9	12	2	US-08-911-092-21	Sequence 21, Appl	680	5	2.9	14	3	US-08-567-357A-13	Sequence 13, Appl
608	5	2.9	12	2	US-08-911-092-22	Sequence 22, Appl	681	5	2.9	14	3	US-08-602-999A-67	Sequence 67, Appl
609	5	2.9	12	2	US-08-911-092-23	Sequence 23, Appl	682	5	2.9	14	3	US-08-729-743A-13	Sequence 13, Appl
610	5	2.9	12	2	US-08-911-092-24	Sequence 24, Appl	683	5	2.9	14	3	US-08-349-498-13	Sequence 13, Appl
611	5	2.9	12	2	US-08-911-092-25	Sequence 25, Appl	684	5	2.9	14	4	US-08-278-865-67	Sequence 67, Appl

685	14	2.9	5	Sequence 67, Appl	758	5	2.9	20	4	US-09-500-124-67	Sequence 1834, Ap
686	14	2.9	5	Sequence 51, Appl	759	5	2.9	20	4	US-08-475-399A-51	Sequence 1835, Ap
687	14	2.9	5	Sequence 38, Appl	760	5	2.9	20	4	US-09-454-204A-38	Sequence 9, Appl
688	14	2.9	5	Sequence 51, Appl	761	5	2.9	20	4	PCT-US93-07545-51	GENERAL INFORMATI
689	14	2.9	5	Sequence 13, Appl	762	5	2.9	20	4	PCT-US95-15463-13	GENERAL INFORMATI
690	14	2.9	5	Sequence 13, Appl	763	5	2.9	20	4	PCT-US95-15923-13	Sequence 63, Appl
691	15	2.9	5	Sequence 27, Appl	764	5	2.9	20	4	US-08-230-047-27	Sequence 4, Appl
692	15	2.9	5	Sequence 30, Appl	765	5	2.9	20	4	US-08-230-047-30	Sequence 11, Appl
693	15	2.9	5	Sequence 31, Appl	766	5	2.9	20	5	US-08-230-047-31	Sequence 11, Appl
694	15	2.9	5	Sequence 39, Appl	767	5	2.9	21	1	US-08-553-257A-39	Sequence 33, Appl
695	15	2.9	5	Sequence 91, Appl	768	5	2.9	21	1	US-08-630-916A-91	Sequence 33, Appl
696	15	2.9	5	Sequence 301, App	769	5	2.9	21	3	US-08-602-999A-301	Sequence 7, Appl
697	15	2.9	5	Sequence 321, App	770	5	2.9	21	3	US-08-602-999A-321	Sequence 3, Appl
698	15	2.9	5	Sequence 322, App	771	5	2.9	21	4	US-08-602-999A-322	Sequence 148, App
699	15	2.9	5	Sequence 355, App	772	5	2.9	21	4	US-08-602-999A-355	Sequence 10, Appl
700	15	2.9	5	Sequence 363, App	773	5	2.9	21	5	US-08-602-999A-363	Sequence 7, Appl
701	15	2.9	5	Sequence 467, App	774	5	2.9	21	6	US-08-602-999A-467	Patent No. 5266328
702	15	2.9	5	Sequence 60, Appl	775	5	2.9	22	2	US-09-406-781-60	Sequence 26, Appl
703	15	2.9	5	Sequence 147, App	776	5	2.9	22	2	US-08-630-915A-147	Sequence 41, Appl
704	15	2.9	5	Sequence 750, App	777	5	2.9	22	3	US-09-149-476-750	Sequence 92, Appl
705	15	2.9	5	Sequence 301, App	778	5	2.9	22	4	US-09-500-124-301	Sequence 41, Appl
706	15	2.9	5	Sequence 321, App	779	5	2.9	22	4	US-09-500-124-321	Sequence 41, Appl
707	15	2.9	5	Sequence 322, App	780	5	2.9	22	6	US-09-500-124-322	Sequence 41, Appl
708	15	2.9	5	Sequence 355, App	781	5	2.9	23	4	US-09-500-124-355	Patent No. 5395759
709	15	2.9	5	Sequence 363, App	782	5	2.9	24	4	US-09-500-124-363	Sequence 25, Appl
710	15	2.9	5	Sequence 467, App	783	5	2.9	25	2	US-09-500-124-467	Sequence 604, App
711	15	2.9	5	Sequence 39, Appl	784	5	2.9	25	2	US-09-500-124-39	Sequence 13, Appl
712	15	2.9	5	Sequence 16, Appl	785	5	2.9	25	2	US-09-880-132-60	Sequence 8, Appl
713	15	2.9	5	Sequence 132, App	786	5	2.9	25	2	US-09-880-132-60	Sequence 20, Appl
714	15	2.9	5	Sequence 3, Appli	787	5	2.9	25	3	US-09-947-372A-16	Sequence 20, Appl
715	15	2.9	5	Sequence 3, Appli	788	5	2.9	25	3	US-09-069-827A-132	Sequence 20, Appl
716	16	2.9	5	Sequence 3, Appli	789	5	2.9	25	3	US-08-470-887A-3	Sequence 20, Appl
717	16	2.9	5	Sequence 3, Appli	790	5	2.9	25	4	US-08-252-508B-3	Sequence 20, Appl
718	16	2.9	5	Sequence 3, Appli	791	5	2.9	25	4	US-09-106-377-3	Sequence 20, Appl
719	16	2.9	5	Sequence 194, App	792	5	2.9	25	4	US-08-602-999A-194	Sequence 20, Appl
720	16	2.9	5	Sequence 213, App	793	5	2.9	26	4	US-08-602-999A-213	Sequence 20, Appl
721	16	2.9	5	Sequence 215, App	794	5	2.9	26	4	US-08-602-999A-215	Sequence 20, Appl
722	16	2.9	5	Sequence 234, App	795	5	2.9	26	4	US-08-602-999A-234	Sequence 20, Appl
723	16	2.9	5	Sequence 194, App	796	5	2.9	27	2	US-08-602-999A-194	Sequence 20, Appl
724	16	2.9	5	Sequence 213, App	797	5	2.9	27	3	US-09-500-124-194	Sequence 20, Appl
725	16	2.9	5	Sequence 215, App	798	5	2.9	27	4	US-09-500-124-213	Sequence 20, Appl
726	16	2.9	5	Sequence 234, App	799	5	2.9	27	4	US-09-500-124-215	Sequence 20, Appl
727	16	2.9	5	Sequence 496, App	800	5	2.9	28	2	US-09-500-124-234	Sequence 20, Appl
728	16	2.9	5	Sequence 496, App	801	5	2.9	30	4	US-08-469-260A-496	Sequence 20, Appl
729	16	2.9	5	Sequence 496, App	802	5	2.9	31	4	US-08-488-446-496	Sequence 20, Appl
730	16	2.9	5	Patent No. 5166057	803	5	2.9	33	3	US-08-467-344A-496	Sequence 26, App
731	17	2.9	5	Sequence 273, App	804	5	2.9	33	3	5166057-16	Sequence 26, App
732	17	2.9	5	Sequence 22, Appl	805	5	2.9	33	3	US-09-428-082B-273	Sequence 28, Appl
733	19	2.9	5	Sequence 11, Appl	806	5	2.9	33	4	US-09-890-650-22	Sequence 15, Appl
734	19	2.9	5	Sequence 11, Appl	807	5	2.9	33	4	US-08-325-553-11	Sequence 1170, Ap
735	19	2.9	5	Sequence 2, Appli	808	5	2.9	33	4	US-08-394-152A-11	Sequence 1171, Ap
736	19	2.9	5	Sequence 17, Appl	809	5	2.9	33	4	US-08-520-933-2	Sequence 1172, Ap
737	19	2.9	5	Sequence 23, Appl	810	5	2.9	33	4	US-09-026-276-17	Sequence 1173, Ap
738	19	2.9	5	Sequence 2, Appli	811	5	2.9	34	1	US-09-053-611-23	Sequence 1174, Ap
739	19	2.9	5	Sequence 11, Appl	812	5	2.9	34	1	US-09-285-040-2	Sequence 18, Appl
740	19	2.9	5	Sequence 17, Appl	813	5	2.9	34	1	US-08-705-477E-11	Sequence 23, Appl
741	20	2.9	5	Sequence 11, Appl	814	5	2.9	34	1	US-09-964-201A-17	Sequence 18, Appl
742	20	2.9	5	Sequence 25, Appl	815	5	2.9	34	2	US-07-678-974D-11	Sequence 23, Appl
743	20	2.9	5	Sequence 4, Appli	816	5	2.9	35	3	US-08-126-016-25	Sequence 23, Appl
744	20	2.9	5	Sequence 4, Appli	817	5	2.9	35	4	US-08-253-751-4	Sequence 18, Appl
745	20	2.9	5	Sequence 11, Appl	818	5	2.9	35	4	US-08-453-925-4	Sequence 49, Appl
746	20	2.9	5	Sequence 16, Appl	819	5	2.9	35	4	US-08-363-276B-11	Sequence 47, Appl
747	20	2.9	5	Sequence 49, Appl	820	5	2.9	35	4	US-08-945-168-16	Sequence 49, Appl
748	20	2.9	5	Sequence 55, Appl	821	5	2.9	36	2	US-08-789-333F-49	Sequence 23, Appl
749	20	2.9	5	Sequence 146, App	822	5	2.9	37	1	US-09-169-015-55	Sequence 4, Appli
750	20	2.9	5	Sequence 148, App	823	5	2.9	37	1	US-08-602-999A-146	Sequence 107, App
751	20	2.9	5	Sequence 11, Appl	824	5	2.9	37	2	US-08-602-999A-148	Sequence 29, Appl
752	20	2.9	5	Sequence 26, Appl	825	5	2.9	37	2	US-08-755-034-11	Sequence 488, App
753	20	2.9	5	Sequence 4, Appli	826	5	2.9	37	3	US-09-240-078-26	Sequence 33, Appl
754	20	2.9	5	Sequence 9, Appli	827	5	2.9	37	3	US-09-142-623-4	Sequence 488, App
755	20	2.9	5	Sequence 146, App	828	5	2.9	37	4	US-08-403-253A-9	Sequence 28, Appl
756	20	2.9	5	Sequence 148, App	829	5	2.9	37	4	US-09-500-124-146	Sequence 488, App
757	20	2.9	5	Sequence 49, Appl	830	5	2.9	37	4	US-09-500-124-148	Sequence 55, Appl
									4	US-08-787-738B-49	Sequence 28, Appl



831	5	2.9	37	4	US-09-454-533-33	Sequence 33, Appl	904	5	2.9	60	4	US-09-638-715-17	Sequence 17, Appl
832	5	2.9	38	3	US-08-630-916A-34	Sequence 34, Appl	905	5	2.9	60	4	US-10-060-509-17	Sequence 17, Appl
833	5	2.9	38	3	US-08-630-916A-35	Sequence 35, Appl	906	5	2.9	60	4	US-09-621-976-5998	Sequence 5998, Ap
834	5	2.9	39	2	US-08-023-980B-34	Sequence 34, Appl	907	5	2.9	60	4	US-09-845-917A-42	Sequence 42, Appl
835	5	2.9	39	2	US-08-486-953A-29	Sequence 29, Appl	908	5	2.9	60	4	US-10-060-506-17	Sequence 17, Appl
836	5	2.9	40	2	US-08-530-569B-7	Sequence 7, Appli	909	5	2.9	61	4	US-09-147-857-5	Sequence 5, Appli
837	5	2.9	40	3	US-08-959-148-2	Sequence 2, Appli	910	5	2.9	61	4	US-09-489-039A-8378	Sequence 8378, Ap
838	5	2.9	40	3	US-09-326-039-7	Sequence 7, Appli	911	5	2.9	62	3	US-09-320-095-2	Sequence 2, Appli
839	5	2.9	42	2	US-08-761-248B-10	Sequence 10, Appl	912	5	2.9	62	3	US-09-320-095-4	Sequence 4, Appli
840	5	2.9	43	4	US-09-144-428-68	Sequence 68, Appl	913	5	2.9	62	3	US-09-523-487-2	Sequence 2, Appli
841	5	2.9	43	4	US-09-148-545-141	Sequence 141, App	914	5	2.9	62	3	US-09-523-487-4	Sequence 4, Appli
842	5	2.9	44	1	US-08-050-319B-37	Sequence 37, Appl	915	5	2.9	62	4	US-08-630-915A-138	Sequence 138, App
843	5	2.9	44	2	US-08-465-982-37	Sequence 37, Appl	916	5	2.9	62	4	US-08-630-915A-218	Sequence 218, App
844	5	2.9	45	3	US-08-965-903B-18	Sequence 18, Appl	917	5	2.9	62	4	US-09-227-357-543	Sequence 543, App
845	5	2.9	45	3	US-08-974-549A-178	Sequence 178, App	918	5	2.9	62	4	US-09-227-357-543	Sequence 543, App
846	5	2.9	45	4	US-09-149-476-641	Sequence 641, App	919	5	2.9	62	4	US-09-134-001C-4040	Sequence 4040, Ap
847	5	2.9	45	4	US-09-402-181B-178	Sequence 178, App	920	5	2.9	62	4	US-09-252-991A-17142	Sequence 17142, A
848	5	2.9	45	4	US-09-721-456-178	Sequence 178, App	921	5	2.9	62	4	US-09-328-352-5365	Sequence 5365, Ap
849	5	2.9	46	2	US-08-726-306A-110	Sequence 110, App	922	5	2.9	62	4	US-09-621-976-5349	Sequence 5349, Ap
850	5	2.9	46	3	US-08-975-600-2	Sequence 2, Appli	923	5	2.9	63	4	US-09-050-010-11	Sequence 11, Appl
851	5	2.9	47	4	US-08-973-131-65	Sequence 65, Appl	924	5	2.9	63	4	US-09-328-352-5621	Sequence 5621, Ap
852	5	2.9	49	6	5252477-4	Patent No. 5252477	925	5	2.9	63	4	US-09-543-681A-4413	Sequence 4413, Ap
853	5	2.9	50	3	US-09-053-197A-53	Sequence 53, Appl	926	5	2.9	63	4	US-09-489-039A-10108	Sequence 10108, A
854	5	2.9	50	4	US-09-085-761A-58	Sequence 58, Appl	927	5	2.9	63	4	US-09-540-236-2004	Sequence 2004, Ap
855	5	2.9	50	4	US-09-621-976-6070	Sequence 6070, Ap	928	5	2.9	64	4	US-09-540-236-2705	Sequence 2705, Ap
856	5	2.9	51	3	US-09-345-468-10	Sequence 10, Appl	929	5	2.9	64	4	US-09-690-454-49	Sequence 49, Appl
857	5	2.9	51	4	US-09-414-453A-10	Sequence 10, Appl	930	5	2.9	64	4	US-09-328-352-5646	Sequence 5646, Ap
858	5	2.9	51	4	US-08-469-260A-347	Sequence 347, App	931	5	2.9	64	4	US-09-621-976-5995	Sequence 5995, Ap
859	5	2.9	51	4	US-08-488-446-347	Sequence 347, App	932	5	2.9	64	4	US-09-643-657-17	Sequence 17, Appl
860	5	2.9	51	4	US-08-467-344A-347	Sequence 347, App	933	5	2.9	65	1	US-08-691-641-11	Sequence 11, Appl
861	5	2.9	52	2	US-08-301-915-5	Sequence 5, Appli	934	5	2.9	65	4	US-09-461-325-452	Sequence 452, App
862	5	2.9	52	3	US-08-524-694A-5	Sequence 5, Appli	935	5	2.9	65	4	US-09-252-991A-23443	Sequence 23443, A
863	5	2.9	52	3	US-08-824-800D-12	Sequence 12, Appl	936	5	2.9	65	4	US-09-252-991A-28764	Sequence 28764, A
864	5	2.9	52	3	US-09-117-217-10	Sequence 10, Appl	937	5	2.9	65	4	US-09-418-710-68	Sequence 68, Appl
865	5	2.9	52	3	US-09-588-751-12	Sequence 12, Appl	938	5	2.9	65	4	US-09-489-039A-7695	Sequence 7695, Ap
866	5	2.9	52	4	US-09-454-156A-5	Sequence 5, Appli	939	5	2.9	65	4	US-09-489-039A-10272	Sequence 10272, A
867	5	2.9	52	4	US-09-735-487-10	Sequence 10, Appl	940	5	2.9	66	4	US-09-328-352-6392	Sequence 6392, Ap
868	5	2.9	52	5	PCT-US95-06266-128	Sequence 128, App	941	5	2.9	67	4	US-09-527-236A-25	Sequence 25, Appl
869	5	2.9	53	2	US-08-726-306A-144	Sequence 144, App	942	5	2.9	67	4	US-09-331-254-2	Sequence 2, Appli
870	5	2.9	53	4	US-09-105-470B-13	Sequence 13, Appl	943	5	2.9	67	4	US-09-756-854-25	Sequence 25, Appl
871	5	2.9	53	4	US-09-621-976-7316	Sequence 7316, Ap	944	5	2.9	68	3	US-08-833-553-9	Sequence 9, Appli
872	5	2.9	54	3	US-08-824-800D-11	Sequence 11, Appl	945	5	2.9	68	3	US-09-418-222-9	Sequence 9, Appli
873	5	2.9	54	3	US-09-588-751-11	Sequence 11, Appl	946	5	2.9	68	4	US-09-489-039A-12470	Sequence 12470, A
874	5	2.9	54	4	US-09-434-840-76	Sequence 76, Appl	947	5	2.9	68	4	US-09-621-976-7121	Sequence 7121, Ap
875	5	2.9	54	4	US-09-621-976-6906	Sequence 6906, Ap	948	5	2.9	69	3	US-08-818-112-78	Sequence 78, Appl
876	5	2.9	55	2	US-08-102-385G-17	Sequence 17, Appl	949	5	2.9	69	4	US-08-818-111-79	Sequence 79, Appl
877	5	2.9	55	4	US-09-144-428-67	Sequence 67, Appl	950	5	2.9	69	4	US-09-056-556-78	Sequence 78, Appl
878	5	2.9	55	4	US-09-144-428-69	Sequence 69, Appl	951	5	2.9	69	4	US-09-072-596-79	Sequence 79, Appl
879	5	2.9	55	4	US-09-621-976-5065	Sequence 5065, Ap	952	5	2.9	69	4	US-09-072-596-79	Sequence 79, Appl
880	5	2.9	55	4	US-09-621-976-5659	Sequence 5659, Ap	953	5	2.9	69	4	US-09-780-016-4	Sequence 4, Appli
881	5	2.9	56	2	US-08-323-449B-3	Sequence 3, Appli	954	5	2.9	69	4	US-09-072-967-78	Sequence 78, Appl
882	5	2.9	56	2	US-08-485-981-3	Sequence 3, Appli	955	5	2.9	69	4	US-09-489-039A-9247	Sequence 9247, Ap
883	5	2.9	56	2	US-08-867-087B-4	Sequence 4, Appli	956	5	2.9	70	3	US-09-367-953B-51	Sequence 51, Appl
884	5	2.9	56	2	US-08-867-087B-26	Sequence 26, Appl	957	5	2.9	70	4	US-09-903-814A-8	Sequence 8, Appli
885	5	2.9	56	5	PCT-US92-08558-8	Sequence 8, Appli	958	5	2.9	71	4	US-09-252-991A-20055	Sequence 20055, A
886	5	2.9	57	1	US-08-379-437-1	Sequence 1, Appli	959	5	2.9	71	4	US-09-673-395A-282	Sequence 31779, A
887	5	2.9	57	1	US-08-379-437-2	Sequence 2, Appli	960	5	2.9	71	4	US-09-252-991A-31779	Sequence 282, App
888	5	2.9	57	1	US-08-379-437-4	Sequence 4, Appli	961	5	2.9	71	4	US-09-621-976-6399	Sequence 6399, Ap
889	5	2.9	57	1	US-08-379-437-6	Sequence 6, Appli	962	5	2.9	71	4	US-09-621-976-6566	Sequence 6566, Ap
890	5	2.9	57	1	US-08-379-437-8	Sequence 8, Appli	963	5	2.9	72	1	US-07-956-700B-75	Sequence 75, Appl
891	5	2.9	57	2	US-08-680-326-115	Sequence 115, App	964	5	2.9	72	1	US-08-476-537-75	Sequence 75, Appl
892	5	2.9	57	2	US-08-680-326-120	Sequence 120, App	965	5	2.9	72	1	US-08-485-607-75	Sequence 75, Appl
893	5	2.9	57	2	US-08-680-326-121	Sequence 121, App	966	5	2.9	72	2	US-08-475-879-75	Sequence 75, Appl
894	5	2.9	57	2	US-08-680-326-122	Sequence 122, App	967	5	2.9	72	4	US-09-433-043B-75	Sequence 75, Appl
895	5	2.9	57	2	US-08-680-326-123	Sequence 123, App	968	5	2.9	72	4	US-09-107-532A-3838	Sequence 3838, Ap
896	5	2.9	57	4	US-09-345-293-6	Sequence 6, Appli	969	5	2.9	72	4	US-09-489-039A-14236	Sequence 14236, A
897	5	2.9	57	4	US-09-904-615-84	Sequence 84, Appl	970	5	2.9	73	4	US-09-311-731A-296	Sequence 6919, Ap
898	5	2.9	57	4	US-09-621-976-4483	Sequence 4483, Ap	971	5	2.9	73	4	US-08-311-731A-296	Sequence 296, App
899	5	2.9	58	1	US-08-465-167A-1	Sequence 1, Appli	972	5	2.9	73	4	US-09-713-550-202	Sequence 202, App
900	5	2.9	58	4	US-08-627-820-1	Sequence 1, Appli	973	5	2.9	73	4	US-09-134-000C-6109	Sequence 6109, Ap
901	5	2.9	59	1	US-08-524-677-16	Sequence 16, Appl	974	5	2.9	74	3	US-09-621-976-7132	Sequence 7132, Ap
902	5	2.9	59	4	US-09-418-710-67	Sequence 67, Appl	975	5	2.9	74	3	US-08-866-545-6	Sequence 6, Appli
903	5	2.9	60	4	US-09-288-143-89	Sequence 89, Appl	976	5	2.9	74	4	US-08-866-545-7	Sequence 7, Appli
												US-08-469-260A-498	Sequence 498, App

```
977      5      2.9      74      4      US-09-252-991A-22657      Sequence 22657, A
978      5      2.9      74      4      US-08-488-446-498      Sequence 498, App
979      5      2.9      74      4      US-08-467-344A-498      Sequence 498, App
980      5      2.9      74      4      US-09-489-039A-10803      Sequence 10803, A
981      5      2.9      74      4      US-09-627-775-6          Sequence 6, Appli
982      5      2.9      74      4      US-09-627-775-7          Sequence 7, Appli
983      5      2.9      75      4      US-09-134-001C-3025      Sequence 3025, Ap
984      5      2.9      75      4      US-08-469-260A-407      Sequence 407, App
985      5      2.9      75      4      US-08-469-260A-478      Sequence 478, App
986      5      2.9      75      4      US-09-252-991A-19131      Sequence 19131, A
987      5      2.9      75      4      US-08-488-446-407      Sequence 407, App
988      5      2.9      75      4      US-08-488-446-478      Sequence 478, App
989      5      2.9      75      4      US-09-732-210-833      Sequence 833, App
990      5      2.9      75      4      US-08-467-344A-407      Sequence 407, App
991      5      2.9      75      4      US-08-467-344A-478      Sequence 478, App
992      5      2.9      75      4      US-09-543-681A-6642      Sequence 6642, Ap
993      5      2.9      75      4      US-09-621-976-7393      Sequence 7393, Ap
994      5      2.9      76      1      US-08-519-777-22        Sequence 22, Appl
995      5      2.9      76      1      US-08-742-035-22        Sequence 22, Appl
996      5      2.9      76      2      US-08-777-019-22        Sequence 22, Appl
997      5      2.9      76      2      US-08-777-143-22        Sequence 22, Appl
998      5      2.9      76      3      US-08-775-414-22        Sequence 22, Appl
999      5      2.9      76      3      US-08-931-858E-22        Sequence 22, Appl
1000     5      2.9      76      3      US-08-981-739-22        Sequence 22, Appl
```

## ALIGNMENTS

```
RESULT 1
US-09-489-039A-11508
; Sequence 11508, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11508
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11508
```

```
Query Match      4.6%; Score 8; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
QY      118 ARTASARP 125
      |||||
Db      57 ARTASARP 64
```

```
RESULT 2
US-09-134-000C-4266
; Sequence 4266, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 4266
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4266

Query Match      4.6%; Score 8; DB 4; Length 181;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      29 ALPTEAAL 36
      |||||
Db      59 ALPTEAAL 66

RESULT 3
US-09-071-035-84
; Sequence 84, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 526 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-84
```

```
Query Match      4.6%; Score 8; DB 4; Length 526;
Best Local Similarity 100.0%; Pred. No. 12;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
QY      29 ALPTEAAL 36
      |||||
Db      404 ALPTEAAL 411
```

```
RESULT 4
US-09-071-035-82
; Sequence 82, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
```





US-08-278-865-93

Query Match 4.0%; Score 7; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLPP 112  
|||||||  
Db 2 STRPLPP 8

RESULT 7

US-09-500-124-93  
; Sequence 93, Application US/09500124  
; Patent No. 6432920

; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/500,124  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/602,999  
; FILING DATE: 16-FEB-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-202  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 93:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide

US-09-500-124-93  
Query Match 4.0%; Score 7; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLPP 112  
|||||||  
Db 2 STRPLPP 8

RESULT 8

US-08-602-999A-60  
; Sequence 60, Application US/08602999A

; Patent No. 6184205

; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/602,999A  
; FILING DATE: 16-FEB-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-202  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide

US-08-602-999A-60  
Query Match 4.0%; Score 7; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLPP 112  
|||||||  
Db 16 STRPLPP 22

RESULT 9

US-08-278-865-60  
; Sequence 60, Application US/08278865  
; Patent No. 6303574

; GENERAL INFORMATION:  
; APPLICANT: KAY, BRIAN K.  
; APPLICANT: SPARKS, ANDREW B.  
; APPLICANT: THORN, JUDITH M.  
; APPLICANT: QUILLIAM, LAWRENCE A.  
; APPLICANT: DER, CHANNING J.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.

US-08-278-865-60  
Query Match 4.0%; Score 7; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



/ ZIP: 22202  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ APPLICANT: Pennie & Edmonds  
/ FILING DATE:  
/ CLASSIFICATION: 514  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Villacorta, Gilberto M.  
/ REGISTRATION NUMBER: 34,038  
/ REFERENCE/DOCKET NUMBER: 4980-007-0  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (703) 413-3000  
/ TELEFAX: (703) 413-2220  
/ TELEX: 248855 OPAT UR  
/ INFORMATION FOR SEQ ID NO: 60:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 31 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: unknown  
/ MOLECULE TYPE: peptide  
US-08-278-865-60

Query Match 4.0%; Score 7; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLPP 112  
|||  
Db 16 STRPLPP 22

## RESULT 10

US-09-500-124-60  
/ Sequence 60, Application US/09500124  
/ Patent No. 6432920

## / GENERAL INFORMATION:

/ APPLICANT: SPARKS, Andrew B.  
/ APPLICANT: KAY, Brian K.  
/ APPLICANT: THORN, Judith M.  
/ APPLICANT: QUILLIAM, Lawrence A.  
/ APPLICANT: DER, Channing J.  
/ APPLICANT: FOWLKES, Dana M.  
/ APPLICANT: RIDER, James E.  
/ TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
/ TITLE OF INVENTION: ISOLATING AND USING SAME  
/ NUMBER OF SEQUENCES: 467

## / CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Pennie & Edmonds  
/ STREET: 1155 Avenue of the Americas  
/ CITY: New York  
/ STATE: New York  
/ COUNTRY: U.S.A.  
/ ZIP: 10036-2711

## / COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICANT: Pennie & Edmonds  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 08/602,999  
/ FILING DATE: 16-FEB-1996  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Misrock, S. Leslie  
/ REGISTRATION NUMBER: 18,872  
/ REFERENCE/DOCKET NUMBER: 1101-202

/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (212) 790-9090  
/ TELEFAX: (212) 869-9741/8864  
/ TELEX: 66141 PENNIE  
/ INFORMATION FOR SEQ ID NO: 60:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 31 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: unknown  
/ MOLECULE TYPE: peptide  
US-09-500-124-60

Query Match 4.0%; Score 7; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLPP 112  
|||  
Db 16 STRPLPP 22

## RESULT 11

US-09-252-991A-23661  
/ Sequence 23661, Application US/09252991A  
/ Patent No. 6551795

## / GENERAL INFORMATION:

/ APPLICANT: Marc J. Rubenfield et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
/ FILE REFERENCE: 107196.136

/ CURRENT APPLICATION NUMBER: US/09/252,991A

/ CURRENT FILING DATE: 1999-02-18

/ PRIOR APPLICATION NUMBER: US 60/074,788

/ PRIOR FILING DATE: 1998-02-18

/ PRIOR APPLICATION NUMBER: US 60/094,190

/ PRIOR FILING DATE: 1998-07-27

/ NUMBER OF SEQ ID NOS: 33142

/ SEQ ID NO 23661

/ LENGTH: 139

/ TYPE: PRT

/ ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23661

Query Match 4.0%; Score 7; DB 4; Length 139;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 SSTRPLP 111  
|||  
Db 76 SSTRPLP 82

## RESULT 12

US-09-252-991A-26927  
/ Sequence 26927, Application US/09252991A  
/ Patent No. 6551795

## / GENERAL INFORMATION:

/ APPLICANT: Marc J. Rubenfield et al.

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

/ FILE REFERENCE: 107196.136

/ CURRENT APPLICATION NUMBER: US/09/252,991A

/ CURRENT FILING DATE: 1999-02-18

/ PRIOR APPLICATION NUMBER: US 60/074,788

/ PRIOR FILING DATE: 1998-02-18

/ PRIOR APPLICATION NUMBER: US 60/094,190

/ PRIOR FILING DATE: 1998-07-27

/ NUMBER OF SEQ ID NOS: 33142

/ SEQ ID NO 26927

/ LENGTH: 171

/ TYPE: PRT

/ ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26927

Query Match 4.0%; Score 7; DB 4; Length 171;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 120 TASARPP 126  
Db 99 TASARPP 105  
  
RESULT 13  
US-09-252-991A-18104  
; Sequence 18104, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18104  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18104

Query Match 4.0%; Score 7; DB 4; Length 177;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 119 RTASARP 125  
Db 134 RTASARP 140  
  
RESULT 14  
US-09-252-991A-21245  
; Sequence 21245, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 21245  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-21245

Query Match 4.0%; Score 7; DB 4; Length 192;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 122 SARPPTS 128  
Db 150 SARPPTS 156  
  
RESULT 15

US-09-252-991A-27353  
; Sequence 27353, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27353  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27353

Query Match 4.0%; Score 7; DB 4; Length 195;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 117 CARTASA 123  
Db 94 CARTASA 100

RESULT 16  
US-09-252-991A-27371  
; Sequence 27371, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27371  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27371

Query Match 4.0%; Score 7; DB 4; Length 236;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 104 TSSTRPL 110  
Db 208 TSSTRPL 214

RESULT 17  
US-08-494-907-6  
; Sequence 6, Application US/08494907  
; Patent No. 5955298  
; GENERAL INFORMATION:  
; APPLICANT: Thomashow, Linda S  
; APPLICANT: Bangera, Mahalaxmi  
; APPLICANT: Weller, David M  
; APPLICANT: Cook, R. James  
; TITLE OF INVENTION: Sequences for Production of  
; TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods  
; NUMBER OF SEQUENCES: 20



```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,907
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0009.95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-494-907-6

Query Match          4.0%; Score 7; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      53 LRPDYAL 59
      |||||
Db      254 LRPDYAL 260

RESULT 18
PCT-US96-10986-6
; Sequence 6, Application PC/TUS9610986
; GENERAL INFORMATION:
; TITLE OF INVENTION: Sequences for Production of
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan A. Pendorf, DOMINIK & STEIN
; STREET: 600 N. West Shore Boulevard, Suite 1000
; CITY: Tampa
; STATE: FL
; COUNTRY: USA
; ZIP: 33609
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10986
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pendorf, Stephan A.
; REGISTRATION NUMBER: 32665
; REFERENCE/DOCKET NUMBER: A700.320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (813) 289-2966
; TELEFAX: (813) 289-2967
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264 amino acids
```

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-10986-6

Query Match          4.0%; Score 7; DB 5; Length 264;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      53 LRPDYAL 59
      |||||
Db      254 LRPDYAL 260

RESULT 19
US-09-252-991A-26018
; Sequence 26018, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26018
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-26018
```

```

Query Match          4.0%; Score 7; DB 4; Length 363;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      122 SARPPTS 128
      |||||
Db      183 SARPPTS 189

RESULT 20
US-09-711-164-393
; Sequence 393, Application US/097111164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THEREOF
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 393
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-711-164-393

Query Match          4.0%; Score 7; DB 4; Length 385;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      73 DALFDLL 79
      |||||
```

Db 369 DALFDLL 375

RESULT 21

US-09-252-991A-22346

; Sequence 22346, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22346

; LENGTH: 444

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22346

Query Match 4.0%; Score 7; DB 4; Length 444;

Best Local Similarity 100.0%; Pred. No. 96;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 RTASARP 125

Db 417 RTASARP 423

RESULT 22

US-09-252-991A-22084

; Sequence 22084, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22084

; LENGTH: 448

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22084

Query Match 4.0%; Score 7; DB 4; Length 448;

Best Local Similarity 100.0%; Pred. No. 97;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 VRAGLLR 54

Db 344 VRAGLLR 350

RESULT 23

US-09-252-991A-28286

; Sequence 28286, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28286

; LENGTH: 469

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28286

Query Match 4.0%; Score 7; DB 4; Length 469;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 RAGLLRP 55

Db 6 RAGLLRP 12

RESULT 24

US-08-591-685-9

; Sequence 9, Application US/08591685

; Patent No. 6083733

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Thermostable xylanases

; NUMBER OF SEQUENCES: 13

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/591,685

; FILING DATE:

; CLASSIFICATION: 435

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 544 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-591-685-9

Query Match 4.0%; Score 7; DB 3; Length 544;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 SVYTSST 107

Db 325 SVYTSST 331

RESULT 25

US-09-252-991A-27527

; Sequence 27527, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: 107196.136

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788



; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27527  
; LENGTH: 760  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27527

Query Match 4.0%; Score 7; DB 4; Length 760;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ALLGHRQ 64  
|||||  
Db 669 ALLGHRQ 675

## RESULT 26

US-09-252-991A-28773  
; Sequence 28773, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28773

; LENGTH: 1007

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28773

Query Match 4.0%; Score 7; DB 4; Length 1007;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AALRTVR 40  
|||||  
Db 879 AALRTVR 885

## RESULT 27

US-08-602-999A-82  
; Sequence 82, Application US/08602999A  
; Patent No. 6184205  
; GENERAL INFORMATION:

; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.

; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/602,999A  
; FILING DATE: 16-FEB-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-202  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 82:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-08-602-999A-82

Query Match 3.4%; Score 6; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLP 111  
|||||

Db 2 STRPLP 7

## RESULT 28

US-08-602-999A-85

; Sequence 85, Application US/08602999A

; Patent No. 6184205

; GENERAL INFORMATION:

; APPLICANT: SPARKS, Andrew B.

; APPLICANT: KAY, Brian K.

; APPLICANT: THORN, Judith M.

; APPLICANT: QUILLIAM, Lawrence A.

; APPLICANT: DER, Channing J.

; APPLICANT: FOWLKES, Dana M.

; APPLICANT: RIDER, James E.

; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF

; TITLE OF INVENTION: ISOLATING AND USING SAME

; NUMBER OF SEQUENCES: 467

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/602,999A

; FILING DATE: 16-FEB-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 1101-202

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 85:

SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-602-999A-85

Query Match 3.4%; Score 6; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLP 111  
Db 2 STRPLP 7

RESULT 29

US-08-602-999A-88  
Sequence 88, Application US/08602999A  
Patent No. 6184205

GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: KAY, Brian K.  
APPLICANT: THORN, Judith M.  
APPLICANT: QUILLIAM, Lawrence A.  
APPLICANT: DER, Channing J.  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: RIDER, James E.  
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,999A  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-602-999A-88

Query Match 3.4%; Score 6; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLP 111  
Db 2 STRPLP 7

RESULT 30

US-08-278-865-82  
Sequence 82, Application US/08278865  
Patent No. 6303574  
GENERAL INFORMATION:

APPLICANT: KAY, BRIAN K.  
APPLICANT: SPARKS, ANDREW B.  
APPLICANT: THORN, JUDITH M.  
APPLICANT: QUILLIAM, LAWRENCE A.  
APPLICANT: DER, CHANNING J.  
TITLE OF INVENTION: SRC SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/278,865  
FILING DATE:

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Villacorta, Gilberto M.  
REGISTRATION NUMBER: 34,038  
REFERENCE/DOCKET NUMBER: 4980-007-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 82:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-278-865-82

Query Match 3.4%; Score 6; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLP 111  
Db 2 STRPLP 7

RESULT 31

US-08-278-865-85  
Sequence 85, Application US/08278865  
Patent No. 6303574  
GENERAL INFORMATION:

APPLICANT: KAY, BRIAN K.  
APPLICANT: SPARKS, ANDREW B.  
APPLICANT: THORN, JUDITH M.  
APPLICANT: QUILLIAM, LAWRENCE A.  
APPLICANT: DER, CHANNING J.  
TITLE OF INVENTION: SRC SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia



RA Terao K., Sugano S.;  
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA  
RT libraries."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB071077; BAB64470.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 246 AA; 26240 MW; 7EE3F9DEEBAPC171 CRC64;  
  
Query Match 4.0%; Score 7; DB 6; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 107 TRPLPPA 113  
|||  
Db 195 TRPLPPA 201  
|||  
  
RESULT 38  
Q9ANV7 PRELIMINARY; PRT; 251 AA.  
AC Q9ANV7;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE YmdA (Fragment).  
GN YMDA.  
OS Lactobacillus delbrueckii.  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OC Lactobacillus.  
OX NCBI\_TaxID=1584;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC11842;  
RA van de Guchte M., Dervyn R., Ehrlich S.D., Maguin E.;  
RT "L. bulgaricus ymdA - rf2 region."  
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF320250; AAK00328.1; -.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR InterPro; IPR006674; HD.  
DR InterPro; IPR003607; Met\_phosphohydro.  
DR InterPro; IPR006675; Unchar\_HDIG.  
DR Pfam; PF01966; HD; 1.  
DR SMART; SM00471; HDC; 1.  
DR TIGRFAMs; TIGR00277; HDIG; 1.  
FT NON TER 1  
SQ SEQUENCE 251 AA; 27719 MW; FF3A99202A09C3E4 CRC64;  
  
Query Match 4.0%; Score 7; DB 2; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 47 AVRAGLL 53  
|||  
Db 92 AVRAGLL 98  
|||  
  
RESULT 39  
Q8IT50 PRELIMINARY; PRT; 262 AA.  
AC Q8IT50;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Fibroin light chain.  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Choi K.;  
RT "Fibroin light chain gene of BaekOk-Jam, Bombyx mori.";

RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF541967; AAN63945.1; -.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR InterPro; IPR008660; L-fibroin.  
DR Pfam; PF05849; L-fibroin; 1.  
SQ SEQUENCE 262 AA; 27652 MW; DC055F5CD9549757 CRC64;  
  
Query Match 4.0%; Score 7; DB 5; Length 262;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 153 VTSAYAA 159  
|||  
Db 11 VTSAYAA 17  
|||  
  
RESULT 40  
O50552 PRELIMINARY; PRT; 264 AA.  
ID O50552;  
AC O50552;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CycB.  
GN CYCB.  
OS Thermus thermophilus.  
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;  
OC Thermus.  
OX NCBI\_TaxID=274;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HB8;  
RA Keightley A., Mather M.W., Fee J.A.;  
RT "Molecular cloning, sequence, and expression of cytochrome c552 from  
RT Thermus thermophilus."  
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
DR EMBL; M93437; AAB88581.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR003439; ABC\_transporter.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
KW ATP-binding; Transport.  
SQ SEQUENCE 264 AA; 29264 MW; 5261060B9CB65138 CRC64;  
  
Query Match 4.0%; Score 7; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 50 AGLLRPD 56  
|||  
Db 71 AGLLRPD 77  
|||  
  
Search completed: May 18, 2004, 16:21:50  
Job time : 72 secs

ID O02210 PRELIMINARY; PRT; 219 AA.  
AC O02210;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE C01A2.6 protein.  
GN C01A2.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kershaw J.K.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL; Z81029; CAB02701.1; -.  
DR PIR; T18797; T18797.  
DR WormPep; C01A2.6; CE07788.  
SQ SEQUENCE 219 AA; 25378 MW; E8C585E206275853 CRC64;  
  
Query Match 4.0%; Score 7; DB 5; Length 219;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 104 TSSTRPL 110  
Db 31 TSSTRPL 37  
  
RESULT 35  
Q8VRF5  
ID Q8VRF5 PRELIMINARY; PRT; 220 AA.  
AC Q8VRF5;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ModC (Molybdenum ABC transporter ATP-binding protein).  
GN MODC OR BLR8162.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tresierra-Ayala A.B., Delgado M.J., Bednar E.J.;  
RT "Characterization and regulation of the Bradyrhizobium japonicum  
modABC genes coding for the molybdenum transport system.";  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
Bradyrhizobium japonicum USDA110.";  
RL DNA Res. 9:189-197(2002).  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
DR EMBL; AF446208; AAL38176.1; -.  
DR EMBL; AP005964; BAC53427.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.

DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR003439; ABC\_transporter.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR SMART; SMO0382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
KW ATP-binding; Transport; Complete proteome.  
SQ SEQUENCE 220 AA; 24264 MW; 553C1F2845103630 CRC64;  
  
Query Match 4.0%; Score 7; DB 16; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 50 AGLLRPD 56  
Db 44 AGLLRPD 50  
  
RESULT 36  
Q7V804  
ID Q7V804 PRELIMINARY; PRT; 242 AA.  
AC Q7V804;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Possible ABC transporter, ATP binding component.  
GN PMT0572.  
OS Prochlorococcus marinus (strain MIT 9313).  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=74547;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22825698; PubMed=12917642;  
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,  
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,  
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
RA Webb E.A., Zinser E.R., Chisholm S.W.;  
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic  
niche differentiation.";  
RL Nature 424:1042-1047(2003).  
DR EMBL; BX572096; CAE20747.1; -.  
KW Complete proteome.  
SQ SEQUENCE 242 AA; 26875 MW; 86344692D2ADB890 CRC64;  
  
Query Match 4.0%; Score 7; DB 16; Length 242;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 51 GLLRPDY 57  
Db 50 GLLRPDY 56  
  
RESULT 37  
Q95LW4  
ID Q95LW4 PRELIMINARY; PRT; 246 AA.  
AC Q95LW4;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,



```
ID Q7Z302 PRELIMINARY; PRT; 173 AA.
AC Q7Z302;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZp686I16132 (Fragment).
GN DKFZp686I16132.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538316; CAD98091.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 173 AA; 19927 MW; 0774F47B1D71E344 CRC64;

Query Match 4.0%; Score 7; DB 4; Length 173;
Best Local Similarity 100.0%; Pred.No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 QPQTQPA 168
Db 162 QPQTQPA 168
|||||

RESULT 32
Q83UG6 PRELIMINARY; PRT; 184 AA.
AC Q83UG6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE HrpL.
GN HrpL.
OS Pseudomonas viridiflava.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=33069;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PNA3.3a, and LP23.1a;
RX MEDLINE=22369509; PubMed=12481991;
RA Jakob K., Goss E.M., Araki H., Van T., Kreitman M., Bergelson J.;
RT "Pseudomonas viridiflava and P. syringae--natural pathogens of
RT Arabidopsis thaliana.";
RL Mol. Plant Microbe Interact. 15:1195-1203 (2002).
DR EMBL; AF508898; AAO73878.1; -.
DR EMBL; AF508899; AAO73879.1; -.
DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
DR GO; GO:0016987; F:sigma factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR009043; RNA_pol_sigma.
DR InterPro; IPR000838; Sigma70_ECF.
DR InterPro; IPR007627; Sigma70_r2.
DR InterPro; IPR007630; Sigma70_r4.
DR Pfam; PF04542; sigma70_r2; 1.
DR Pfam; PF04545; sigma70_r4; 1.
DR PROSITE; PS01063; SIGMA70_ECF; 1.
SQ SEQUENCE 184 AA; 21115 MW; 4F7AB55F877EB2D4 CRC64;

Query Match 4.0%; Score 7; DB 2; Length 184;
Best Local Similarity 100.0%; Pred.No. 98;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GHRQLVR 67
|||||
```

```
Db 119 GHRQLVR 125
|||||

RESULT 33
Q9XA90 PRELIMINARY; PRT; 215 AA.
ID Q9XA90
AC Q9XA90;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative methyltransferase.
GN SCO0835 OR SCF43A.25C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K., Harris D.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL; AL939106; CAB48912.1; -.
DR PIR; T36448; T36448.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000051; SAM_bind.
DR InterPro; IPR008854; TPMT.
DR Pfam; PF05724; TPMT; 1.
KW Methyltransferase; Transferase; Complete proteome.
SQ SEQUENCE 215 AA; 23055 MW; C8DD0383A9D00365 CRC64;

Query Match 4.0%; Score 7; DB 16; Length 215;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 TARRPSV 102
Db 186 TARRPSV 192
|||||

RESULT 34
O02210
```

RP SEQUENCE FROM N.A.  
RC STRAIN=KIM5 / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM.";  
RL J. Bacteriol. 184:4601-4611(2002).  
DR EMBL; AE013648; AAM84067.1; --  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0015450; F:protein translocase activity; IEA.  
DR GO; GO:0009306; P:protein secretion; IEA.  
DR GO; GO:0006605; P:protein targeting; IEA.  
DR InterPro; IPR001901; SecE.  
DR InterPro; IPR005807; SecE\_bac.  
DR Pfam; PF00584; SecE; 1.  
DR PRINTS; PR01650; SECETRNLCASE.  
DR TIGRFAMs; TIGR00964; 3a0501s06; 1.  
DR PROSITE; PS01067; SECE SEC61G; 1.  
SQ SEQUENCE 132 AA; 14305 MW; 503FDD8E42A583FE CRC64;  
  
Query Match 4.0%; Score 7; DB 16; Length 132;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 20 VAIVGNY 26  
|||  
Db 33 VAIVGNY 39  
  
RESULT 28  
O01257  
ID O01257 PRELIMINARY; PRT; 144 AA.  
AC O01257;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE T20D3.6 protein.  
GN T20D3.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lloyd C.R.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL; Z68220; CAA92488.1; --  
DR PIR; T25028; T25028.  
DR WormPep; T20D3.6; CE03674.  
DR InterPro; IPR007667; HIG\_1\_N.  
DR Pfam; PF04588; HIG\_1\_N; 1.  
SQ SEQUENCE 144 AA; 15138 MW; D6339303B565B024 CRC64;  
  
Query Match 4.0%; Score 7; DB 5; Length 144;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 31 PTEAALR 37  
|||  
Db 114 PTEAALR 120  
  
RESULT 29

Q98LQ7  
ID Q98LQ7 PRELIMINARY; PRT; 148 AA.  
AC Q98LQ7;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein ml10920.  
GN ML10920.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP002996; BAB48406.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 148 AA; 16043 MW; 0E5775508A67A79D CRC64;  
  
Query Match 4.0%; Score 7; DB 16; Length 148;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 36 LRTVRDT 42  
|||  
Db 84 LRTVRDT 90  
  
RESULT 30  
P79092  
ID P79092 PRELIMINARY; PRT; 161 AA.  
AC P79092;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Beta casein (Fragment).  
OS Ailurus fulgens (Lesser panda).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Procyonidae; Ailurus.  
OX NCBI\_TaxID=9649;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gatesy J.;  
RT "More DNA support for a Cetacea/Hippopotamidae clade: the blood  
RT clotting protein gene g-fibrinogen.";  
RL Mol. Biol. Evol. 0:0-0(1997).  
DR EMBL; U86646; AAB47432.1; --  
DR InterPro; IPR001588; Casein.  
DR Pfam; PF00363; caseins; 1.  
FT NON\_TER 1  
FT NON\_TER 161  
SQ SEQUENCE 161 AA; 18067 MW; 8E633ED5397C3549 CRC64;  
  
Query Match 4.0%; Score 7; DB 6; Length 161;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 161 AQPQTQP 167  
|||  
Db 88 AQPQTQP 94  
  
RESULT 31  
Q7Z302



Q81X50  
ID Q81X50 PRELIMINARY; PRT; 126 AA.  
AC Q81X50;  
DT 01-JUN-2003 (TReMBLrel. 24, Created)  
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Membrane protein, putative.  
GN BA5393.  
OS Bacillus anthracis (strain Ames).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=198094;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22608414; PubMed=12721629;  
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,  
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,  
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,  
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,  
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,  
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,  
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,  
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,  
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,  
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,  
RA Fraser C.M.;  
RT "The genome sequence of Bacillus anthracis Ames and comparison to  
RT closely related bacteria.";  
RL Nature 423:81-86(2003).  
DR EMBL; AE017040; AAP29052.1; -.  
DR TIGR; BA5393; -.  
DR InterPro; IPR007165; DUF360.  
DR Pfam; PF04020; DUF360; 1.  
KW Complete proteome.  
SQ SEQUENCE 126 AA; 13603 MW; A056B91A63241AAA CRC64;

Query Match 4.0%; Score 7; DB 16; Length 126;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GFGVAIV 23  
| | | | |  
Db 94 GFGVAIV 100

RESULT 25  
Q81516  
ID Q81516 PRELIMINARY; PRT; 126 AA.  
AC Q81516;  
DT 01-JUN-2003 (TReMBLrel. 24, Created)  
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Integral membrane protein.  
GN BC5165.  
OS Bacillus cereus (strain ATCC 14579 / DSM 31).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=226900;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22608415; PubMed=12721630;  
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,  
RA Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,  
RA Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,  
RA Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,  
RA Overbeek R., Kyrpides N.;  
RT "Genome sequence of Bacillus cereus and comparative analysis with  
RT Bacillus anthracis.";  
RL Nature 423:87-91(2003).  
DR EMBL; AE017014; AAP12030.1; -.  
DR InterPro; IPR007165; DUF360.  
DR Pfam; PF04020; DUF360; 1.  
KW Complete proteome.  
SQ SEQUENCE 126 AA; 13603 MW; A056B91A63241AAA CRC64;

Query Match 4.0%; Score 7; DB 16; Length 126;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 17 GFGVAIV 23  
| | | | |  
Db 94 GFGVAIV 100

RESULT 26  
Q8ZAN9  
ID Q8ZAN9 PRELIMINARY; PRT; 127 AA.  
AC Q8ZAN9;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Preprotein translocase SecE subunit.  
GN SECE OR YP03753.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360;  
RA Parkhill J., Wren B.W., Thomson N.R., Titchall R.W., Holden M.T.G.,  
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";  
RL Nature 413:523-527(2001).  
DR EMBL; AJ414158; CAC93221.1; -.  
DR PIR; A10456; A10456.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0015450; F:protein translocase activity; IEA.  
DR GO; GO:0009306; P:protein secretion; IEA.  
DR GO; GO:0006605; P:protein targeting; IEA.  
DR InterPro; IPR001901; SecE.  
DR InterPro; IPR005807; SecE\_bac.  
DR Pfam; PF00584; SecE; 1.  
DR PRINTS; PR01650; SECETRNLCASE.  
DR TIGRFAMs; TIGR00964; 3a0501s06; 1.  
DR PROSITE; PS01067; SECE\_SEC61G; 1.  
KW Complete proteome.  
SQ SEQUENCE 127 AA; 13682 MW; D5E98D0BEA6AA0A9 CRC64;

Query Match 4.0%; Score 7; DB 16; Length 127;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
| | | | |  
Db 28 VAIVGNY 34

RESULT 27  
Q8D1H4  
ID Q8D1H4 PRELIMINARY; PRT; 132 AA.  
AC Q8D1H4;  
DT 01-MAR-2003 (TReMBLrel. 23, Created)  
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Preprotein translocase.  
GN SECE OR Y0478.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]

```
OX  NCBI_TaxID=8255;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Aoki T., Hirono I., Lee J., Iwahori A.;
RT  "Molecular cloning and expression of IL-1b and two types of chemokines
in Japanese flounder, Paralichthys olivaceus.";
RL  Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AB070837; BAB86884.1; -.
DR  GO; GO:0005576; C:extracellular; IEA.
DR  GO; GO:0008009; F:chemokine activity; IEA.
DR  GO; GO:0006955; P:immune response; IEA.
DR  InterPro; IPR001811; Chemokine IL8.
DR  InterPro; IPR001089; CXC_chemkine_sm11.
DR  Pfam; PF00048; IL8; 1.
DR  PRINTS; PR00437; SMALLCYTKXC.
DR  SMART; SM00199; SCY; 1.
SQ  SEQUENCE 98 AA; 10846 MW; 495CA4839A5D8C92 CRC64;

Query Match 4.0%; Score 7; DB 13; Length 98;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AVSLRSL 93
Db 23 AVSLRSL 29
|||||

RESULT 21
Q13862
ID Q13862 PRELIMINARY; PRT; 106 AA.
AC Q13862;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE DNA-binding protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95194400; PubMed=7887923;
RA Luzi P., Strayer D.S.;
RT "DNA binding proteins that amplify surfactant protein B gene
expression: isolation and characterization.";
RL Biochem. Biophys. Res. Commun. 208:153-160(1995).
DR EMBL; L10403; AAA68277.1; -.
DR PIR; I52306; I52306.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0007585; P:respiratory gaseous exchange; TAS.
DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.
KW DNA-binding.
SQ SEQUENCE 106 AA; 11066 MW; 10FE21A2B5685A13 CRC64;

Query Match 4.0%; Score 7; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 FAGHSAG 147
Db 86 FAGHSAG 92
|||||

RESULT 22
Q90Y59
ID Q90Y59 PRELIMINARY; PRT; 109 AA.
AC Q90Y59;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Interleukine-8.
GN IL-8.
```

```
OS Paralichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Paralichthyidae; Paralichthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21435923; PubMed=11551708;
RA Lee E.Y., Park H.H., Kim Y.T., Chung J.K., Choi T.J.;
RT "Cloning and sequence analysis of the interleukin-8 gene from flounder
(Paralichthys olivaceus).";
RL Gene 274:237-243(2001).
DR EMBL; AF216646; AAL05442.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008009; F:chemokine activity; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR001811; Chemokine IL8.
DR InterPro; IPR001089; CXC_chemkine_sm11.
DR Pfam; PF00048; IL8; 1.
DR PRINTS; PR00437; SMALLCYTKXC.
DR SMART; SM00199; SCY; 1.
SQ SEQUENCE 109 AA; 12117 MW; C7CE18986C50A6ED CRC64;

Query Match 4.0%; Score 7; DB 13; Length 109;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AVSLRSL 93
Db 22 AVSLRSL 28
|||||

RESULT 23
Q8YKB7
ID Q8YKB7 PRELIMINARY; PRT; 123 AA.
AC Q8YKB7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Alr7384.
GN ALR7384.
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120alpha.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003601; BAB77142.1; -.
DR PIR; AH2525; AH2525.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 123 AA; 12757 MW; 2FC44D9705E2EF8C CRC64;

Query Match 4.0%; Score 7; DB 16; Length 123;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AALRTVR 40
Db 41 AALRTVR 47
|||||

RESULT 24
```



```
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      160 SAQPQTQP 167
      |||||
Db      407 SAQPQTQP 414

RESULT 17
Q8BTW4
ID      Q8BTW4      PRELIMINARY;      PRT;      667 AA.
AC
DT      01-MAR-2003 (TREMBLrel. 23, Created)
DT      01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Zinc finger protein 282.
GN      ZFP282 OR A149432 OR E430019K09RIK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NOD; TISSUE=Thymus;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
DR      EMBL; AK088531; BAC40406.1; -.
DR      MGD; MGI:2141413; Zfp282.
DR      GO; GO:0005622; C:intracellular; IEA.
DR      GO; GO:0005874; C:microtubule; IEA.
DR      GO; GO:0003676; F:nucleic acid binding; IEA.
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      GO; GO:0007018; P:microtubule-based movement; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR      InterPro; IPR001909; KRAB.
DR      InterPro; IPR000217; Tubulin.
DR      InterPro; IPR007087; Znf_C2H2.
DR      Pfam; PF01352; KRAB; 1.
DR      Pfam; PF00096; zf-C2H2; 5.
DR      ProDom; PD000003; Znf_C2H2; 4.
DR      SMART; SM00349; KRAB; 1.
DR      SMART; SM00355; Znf_C2H2; 5.
DR      PROSITE; PS00805; KRAB; 1.
DR      PROSITE; PS00227; TUBULIN; 1.
DR      PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR      PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.
SQ      SEQUENCE      667 AA; 73593 MW;  A21D5DD1CF12EE59 CRC64;

Query Match      4.6%; Score 8; DB 11; Length 667;
Best Local Similarity      100.0%; Pred. No. 29;
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      160 SAQPQTQP 167
      |||||
Db      407 SAQPQTQP 414

RESULT 18
Q80TG5
ID      Q80TG5      PRELIMINARY;      PRT;      1545 AA.
AC      Q80TG5;
DT      01-JUN-2003 (TREMBLrel. 24, Created)
DT      01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      MKIAA1256 protein (Fragment).
GN      MKIAA1256.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
```

```
[1]
RN      SEQUENCE FROM N.A.
RP      TISSUE=Brain;
RX      MEDLINE=22579291; PubMed=12693553;
RA      Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA      Nakajima D., Nagase T., Ohara O., Koga H.;
RT      "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT      II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT      cDNAs identified by screening of terminal sequences of cDNA clones
RT      randomly sampled from size-fractionated libraries.";
RL      DNA Res. 10:35-48(2003).
DR      EMBL; AK122480; BAC65762.2; -.
FT      NON TER      1
SQ      SEQUENCE      1545 AA; 176155 MW;  024E4F68F9D0EEA9 CRC64;

Query Match      4.6%; Score 8; DB 11; Length 1545;
Best Local Similarity      100.0%; Pred. No. 62;
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      153 VTSAYAAAS 160
      |||||
Db      902 VTSAYAAAS 909

RESULT 19
Q993U5
ID      Q993U5      PRELIMINARY;      PRT;      80 AA.
AC      Q993U5;
DT      01-JUN-2001 (TREMBLrel. 17, Created)
DT      01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      ORF1b polyprotein (Fragment).
OS      Porcine reproductive and respiratory syndrome virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC      Arteriviridae; Arterivirus.
OX      NCBI_TaxID=28344;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=16244B;
RX      MEDLINE=20499111; PubMed=11044133;
RA      Allende R., Laegreid W.W., Kutish G.F., Galeota J.A., Wills R.W.,
RA      Osorio F.A.;
RT      "Porcine Reproductive and Respiratory Syndrome Virus: Description of
RT      Persistence in Individual Pigs upon Experimental Infection.";
RL      J. Virol. 74:10834-10837(2000).
DR      EMBL; AF299415; AAK20545.1; -.
KW      Polyprotein.
FT      NON TER      1
SQ      SEQUENCE      80 AA; 9114 MW;  A63BABD963143B8D CRC64;

Query Match      4.0%; Score 7; DB 12; Length 80;
Best Local Similarity      100.0%; Pred. No. 46;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      107 TRPLPPA 113
      |||||
Db      74 TRPLPPA 80

RESULT 20
Q8QGV8
ID      Q8QGV8      PRELIMINARY;      PRT;      98 AA.
AC      Q8QGV8;
DT      01-JUN-2002 (TREMBLrel. 21, Created)
DT      01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT      01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE      CXC chemokine.
GN      CXC CHEMOKINE.
OS      Paralichthys olivaceus (Flounder).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC      Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC      Pleuronectoidae; Paralichthyidae; Paralichthys.
```

```
OX  NCBI_TaxID=323;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=DC3000;
RA  Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA  Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA  Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA  Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA  White O., Fraser C., Collmer A.;
RT  "Complete sequence of Pseudomonas syringae.";
RL  Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE016872; AAO58057.1; -.
DR  TIGR; PSPTO4611; -.
KW  Complete proteome.
SQ  SEQUENCE 346 AA; 36891 MW; B0C3528561944EB5 CRC64;

      Query Match          4.6%; Score 8; DB 16; Length 346;
      Best Local Similarity 100.0%; Pred. No. 16;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  131 HVYSGNLG 138
Db  276 HVYSGNLG 283
|||||

RESULT 14
Q93TD8      Q93TD8      PRELIMINARY;      PRT;      351 AA.
AC  Q93TD8;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Hypothetical protein.
OS  Pseudomonas syringae pv. maculicola.
OG  Plasmid pFKN.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC  Pseudomonadaceae; Pseudomonas.
OX  NCBI_TaxID=59511;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=M6;
RA  Rohmer L.M., Marchesini P., Kjemtrup S., Chang J.H., Dangl J.L.;
RT  "Complete Nucleotide Sequence of Plasmid pFKN from Pseudomonas
RT  syringae pathovar maculicola Harboring a Pathogenicity Island.";
RL  Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF359557; AAK49552.1; -.
DR  GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW  Hypothetical protein; Plasmid.
SQ  SEQUENCE 351 AA; 37707 MW; 48B43F7F18781A7D CRC64;

      Query Match          4.6%; Score 8; DB 2; Length 351;
      Best Local Similarity 100.0%; Pred. No. 16;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  131 HVYSGNLG 138
Db  281 HVYSGNLG 288
|||||

RESULT 15
Q833Y2      Q833Y2      PRELIMINARY;      PRT;      546 AA.
AC  Q833Y2;
DT  01-JUN-2003 (TrEMBLrel. 24, Created)
DT  01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Pheromone binding protein.
GN  EF1791.
OS  Enterococcus faecalis (Streptococcus faecalis).
OC  Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX  NCBI_TaxID=1351;
RN  [1]
RP  SEQUENCE FROM N.A.
```

```
RC  STRAIN=V583 / ATCC 700802;
RX  MEDLINE=22550857; PubMed=12663927;
RA  Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA  Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA  Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA  Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA  Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA  Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT  "Role of mobile DNA in the evolution of vancomycin-resistant
RT  Enterococcus faecalis.";
RL  Science 299:2071-2074(2003).
DR  EMBL; AE016952; AAC81561.1; -.
DR  TIGR; EF1791; -.
DR  GO; GO:0005215; F:transporter activity; IEA.
DR  GO; GO:0006810; P:transport; IEA.
DR  InterPro; IPR000437; Prok_lipoprot_s.
DR  InterPro; IPR000914; SBP_bac_5.
DR  Pfam; PF00496; SBP_bac_5; 1.
DR  PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW  Complete proteome.
SQ  SEQUENCE 546 AA; 61264 MW; 73737A52ABEF9C60 CRC64;

      Query Match          4.6%; Score 8; DB 16; Length 546;
      Best Local Similarity 100.0%; Pred. No. 24;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  29 ALPTEAAL 36
Db  424 ALPTEAAL 431
|||||

RESULT 16
Q8R3D4      Q8R3D4      PRELIMINARY;      PRT;      657 AA.
AC  Q8R3D4;
DT  01-JUN-2002 (TrEMBLrel. 21, Created)
DT  01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Hypothetical protein.
GN  ZFP282 OR AI449432.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Strausberg R.;
RL  Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC025581; AAK25581.1; -.
DR  MGD; MGI:2141413; Zfp282.
DR  GO; GO:0005622; C:intracellular; IEA.
DR  GO; GO:0005874; C:microtubule; IEA.
DR  GO; GO:0003676; F:nucleic acid binding; IEA.
DR  GO; GO:0005198; F:structural molecule activity; IEA.
DR  GO; GO:0007018; P:microtubule-based movement; IEA.
DR  GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR  InterPro; IPR01909; KRAB.
DR  InterPro; IPR00217; Tubulin.
DR  InterPro; IPR07087; Znf_C2H2.
DR  Pfam; PF01352; KRAB; 1.
DR  Pfam; PF00096; zf-C2H2; 2.
DR  ProDom; PD000003; Znf_C2H2; 1.
DR  SMART; SM00349; KRAB; 1.
DR  SMART; SM00355; Znf_C2H2; 2.
DR  PROSITE; PS50805; KRAB; 1.
DR  PROSITE; PS00227; TUBULIN; 1.
DR  PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR  PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW  Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ  SEQUENCE 657 AA; 72088 MW; C688E4C0FE976BC6 CRC64;

      Query Match          4.6%; Score 8; DB 11; Length 657;
      Best Local Similarity 100.0%; Pred. No. 29;
```



RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003530; AAF49599.1; -.  
DR EMBL; AY069753; AAL39898.1; -.  
DR FlyBase; FBgn004590; Eig71EC.  
DR InterPro; IPR003475; Insect\_Unk.  
DR Pfam; PF02448; L71; 1.  
SQ SEQUENCE 173 AA; 20238 MW; 6C12A1FAE63BBSB2 CRC64;  
  
Query Match 4.6%; Score 8; DB 5; Length 173;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 123 ARPPTSRR 130  
Db |||||||  
102 ARPPTSRR 109  
  
RESULT 11  
Q24073 PRELIMINARY; PRT; 173 AA.  
AC Q24073;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE L71-3.  
GN EIG71EC OR L71-3 OR CG7608.

OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OREGON-R; TISSUE=Salivary gland;  
RA Wright L.G., Chen T., Thummel C.S., Guild G.M.;  
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U24242; AAA65111.1; -.  
DR PIR; S62349; S62349.  
DR FlyBase; FBgn0004590; Eig71EC.  
DR InterPro; IPR003475; Insect\_Unk.  
DR Pfam; PF02448; L71; 1.  
SQ SEQUENCE 173 AA; 20270 MW; EE91F0AAF32FEB47 CRC64;  
  
Query Match 4.6%; Score 8; DB 5; Length 173;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 123 ARPPTSRR 130  
Db |||||||  
102 ARPPTSRR 109  
  
RESULT 12  
Q49675 PRELIMINARY; PRT; 336 AA.  
ID Q49675  
AC Q49675;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE PabB.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96084954; PubMed=7489918;  
RA Doukhan L., Predich M., Nair G., Dussurget O., Mandic-Mulec I.,  
RA Cole S.T., Smith D.R., Smith I.;  
RT "Genomic organization of the mycobacterial sigma gene cluster.";  
RL Gene 165:67-70(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Robison K.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U00012; AAA85925.1; -.  
SQ SEQUENCE 336 AA; 35916 MW; 0CE7CABDB3CBD99F CRC64;  
  
Query Match 4.6%; Score 8; DB 2; Length 336;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 121 ASARPPPTS 128  
Db |||||||  
211 ASARPPPTS 218  
  
RESULT 13  
Q87WE0 PRELIMINARY; PRT; 346 AA.  
ID Q87WE0  
AC Q87WE0;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Conserved domain protein.  
GN PSPTO4611.  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.

```
RESULT 7
Q8VCS0
ID Q8VCS0 PRELIMINARY; PRT; 530 AA.
AC Q8VCS0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Peptidoglycan recognition protein-like (Peptidoglycan recognition
protein L) (EC 3.5.1.28).
GN PGLYRPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Gelius E., Persson C., Karlsson J., Steiner H.K.;
RT "A Mammalian Peptidoglycan Recognition Protein with N-
Acetylmuramoyl-L-Alanine Amidase Activity.";
RT Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC019396; AAH19396.1; -.
DR EMBL; AY282722; AAP22283.1; -.
DR MGD; MGI:1928099; Pglyrp1.
DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro; IPR002502; Amidase_2.
DR InterPro; IPR006619; PGRP.
DR Pfam; PF01510; Amidase_2; 1.
DR SMART; SM00644; Ami2_1.
DR SMART; SM00701; PGRP; 1.
KW Hydrolase.
SQ SEQUENCE 530 AA; 57706 MW; D3BF52597CE5D1F9 CRC64;

Query Match 5.7%; Score 10; DB 11; Length 530;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGHWVGAHT 10
Db 441 RGHWVGAHT 450

RESULT 8
Q23659
ID Q23659 PRELIMINARY; PRT; 161 AA.
AC Q23659;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ZK899.1 protein.
GN ZK899.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
```

```
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z37140; CAA85496.1; -.
DR PIR; T28088; T28088.
DR WormPep; ZK899.1; CE011118.
SQ SEQUENCE 161 AA; 18225 MW; 5F49B3DCCF5FF18A CRC64;

Query Match 5.2%; Score 9; DB 5; Length 161;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 120 TASARPPPTS 128
Db 13 TASARPPPTS 21

RESULT 9
Q9Y8M3
ID Q9Y8M3 PRELIMINARY; PRT; 119 AA.
AC Q9Y8M3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE2609.
GN APE2609.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=103829966;
RA Kwarababayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000064; BAA81627.1; -.
DR PIR; C72496; C72496.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 119 AA; 12825 MW; 991DCE32F4BF22E7 CRC64;

Query Match 4.6%; Score 8; DB 17; Length 119;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 123 ARPPTSRR 130
Db 48 ARPPTSRR 55

RESULT 10
Q9VUS6
ID Q9VUS6 PRELIMINARY; PRT; 173 AA.
AC Q9VUS6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE EIG71EC protein (LP111175p).
GN EIG71EC OR CG7608.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
```



```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Peptidoglycan recognition protein L isoform A.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Sang Y., Ross C.R., Blecha F.;
RT "Characterization of porcine peptidoglycan recognition proteins: gene
RT cloning and regulation on innate immunity.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF541955; AAO41115.1; -.
DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro; IPR002502; Amidase_2.
DR InterPro; IPR006619; PGRP.
DR Pfam; PF01510; Amidase_2; 1.
DR SMART; SM00644; Ami_2; 1.
DR SMART; SM00701; PGRP; 1.
SQ SEQUENCE 252 AA; 28392 MW; 365F34308D2E3560 CRC64;

Query Match 12.6%; Score 22; DB 6; Length 252;
Best Local Similarity 100.0%; Pred. No. 6e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 PDYALLGHRQLVTRDPCGDALF 76
Db 191 PDYALLGHRQLVTRDPCGDALF 212

RESULT 4
Q866Y3
ID Q866Y3 PRELIMINARY; PRT; 598 AA.
AC Q866Y3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Peptidoglycan recognition protein L isoform B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Sang Y., Ross C.R., Blecha F.;
RT "Characterization of porcine peptidoglycan recognition proteins: gene
RT cloning and regulation on innate immunity.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF541956; AAO41116.1; -.
DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro; IPR002502; Amidase_2.
DR InterPro; IPR006619; PGRP.
DR Pfam; PF01510; Amidase_2; 1.
DR SMART; SM00644; Ami_2; 1.
DR SMART; SM00701; PGRP; 1.
SQ SEQUENCE 598 AA; 64593 MW; FCDD237A9F105DDB CRC64;

Query Match 12.6%; Score 22; DB 6; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 PDYALLGHRQLVTRDPCGDALF 76
Db 537 PDYALLGHRQLVTRDPCGDALF 558

RESULT 5
Q9QXZ1
ID Q9QXZ1 PRELIMINARY; PRT; 500 AA.
```

```
AC Q9QXZ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TAGL-beta.
GN PGLYRPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kibardin A.V.;
RT "Molecular cloning of the mouse Tagl-beta gene.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF149838; AAF22234.1; -.
DR MGD; MGI:1928099; Pglyrpl.
DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro; IPR002502; Amidase_2.
DR InterPro; IPR006619; PGRP.
DR Pfam; PF01510; Amidase_2; 1.
DR SMART; SM00644; Ami_2; 1.
DR SMART; SM00701; PGRP; 1.
SQ SEQUENCE 500 AA; 54448 MW; 4548810D215C7887 CRC64;

Query Match 5.7%; Score 10; DB 11; Length 500;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHWVGAHT 10
Db 412 RGWHWVGAHT 421

RESULT 6
Q9QXZ2
ID Q9QXZ2 PRELIMINARY; PRT; 529 AA.
AC Q9QXZ2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TAGL-alpha.
GN PGLYRPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kibardin A.V.;
RT "Molecular cloning of the mouse Tagl-alpha gene.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF149837; AAF22233.1; -.
DR MGD; MGI:1928099; Pglyrpl.
DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro; IPR002502; Amidase_2.
DR InterPro; IPR006619; PGRP.
DR Pfam; PF01510; Amidase_2; 1.
DR SMART; SM00644; Ami_2; 1.
DR SMART; SM00701; PGRP; 1.
SQ SEQUENCE 529 AA; 57603 MW; 82704CCF6C8EC6EA CRC64;

Query Match 5.7%; Score 10; DB 11; Length 529;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGWHWVGAHT 10
Db 441 RGWHWVGAHT 450
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:17:20 ; Search time 39 Seconds  
(without alignments)  
1407.697 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 174  
Sequence: 1 RGHVWGVAHTLGHNSRGFV.....SAYAASQPQTQPACFPSS 174

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 1017041 seqs, 315518202 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL 25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archheap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	174	100.0	634	4	Q96N74	Q96n74 homo sapien
2	87	50.0	576	4	Q96PD5	Q96pd5 homo sapien
3	22	12.6	252	6	Q866Y4	Q866y4 sus scrofa
4	22	12.6	598	6	Q866Y3	Q866y3 sus scrofa
5	10.	5.7	500	11	Q9QXZ1	Q9qxz1 mus musculu
6	10	5.7	529	11	Q9QXZ2	Q9qxz2 mus musculu
7	10	5.7	530	11	Q8VCS0	Q8vcs0 mus musculu
8	9	5.2	161	5	Q23659	Q23659 caenorhabdi
9	8	4.6	119	17	Q9Y8M3	Q9y8m3 aeropyrum p
10	8	4.6	173	5	Q9VUS6	Q9vus6 drosophila
11	8	4.6	173	5	Q24073	Q24073 drosophila
12	8	4.6	336	2	Q49675	Q49675 mycobacteri
13	8	4.6	346	16	Q87WE0	Q87we0 pseudomonas
14	8	4.6	351	2	Q93TD8	Q93td8 pseudomonas
15	8	4.6	546	16	Q833Y2	Q833y2 enterococcu
16	8	4.6	657	11	Q8R3D4	Q8r3d4 mus musculu

90	7	4.0	533	10	Q9LR87	Q9lr87 arabidopsis	163	6	3.4	70	10	Q8LES7	Q8les7 arabidopsis
91	7	4.0	533	16	Q8PHP3	Q8php3 xanthomonas	164	6	3.4	71	10	Q94EB9	Q94eb9 oryza sativ
92	7	4.0	533	16	Q8P694	Q8p694 xanthomonas	165	6	3.4	76	3	Q8X1K7	Q8xlk7 calonectria
93	7	4.0	534	11	Q8BUX7	Q8bux7 mus musculu	166	6	3.4	76	3	Q8X1K6	Q8xlk6 calonectria
94	7	4.0	560	16	Q8YRX9	Q8yrx9 anabaena sp	167	6	3.4	76	10	Q7X721	Q7x721 oryza sativ
95	7	4.0	579	13	Q57397	Q57397 oncorhynch	168	6	3.4	78	16	Q82UL8	Q82ul8 nitrosomona
96	7	4.0	581	16	Q8RBC8	Q8rbcb thermoanaer	169	6	3.4	78	17	Q8TNT8	Q8tnt8 methanosarc
97	7	4.0	594	4	Q8NEV1	Q8nbv1 homo sapien	170	6	3.4	79	2	Q32950	Q32950 mycobacteri
98	7	4.0	598	5	Q8SSN1	Q8ssn1 dictyosteli	171	6	3.4	79	16	Q98FA2	Q98fa2 rhizobium l
99	7	4.0	600	4	Q8N369	Q8n369 homo sapien	172	6	3.4	81	16	Q82UW7	Q82uw7 nitrosomona
100	7	4.0	606	2	Q46606	Q46606 deesulfovibr	173	6	3.4	81	17	Q8TQ12	Q8tql2 methanosarc
101	7	4.0	608	12	Q80MY5	Q80my5 human cytom	174	6	3.4	82	16	Q9ABE1	Q9abe1 caulobacter
102	7	4.0	608	12	Q80MY1	Q80my1 human cytom	175	6	3.4	82	16	Q7UJL9	Q7ujl9 rhodopirell
103	7	4.0	617	3	Q9P6C4	Q9p6c4 neurospora	176	6	3.4	82	17	Q26575	Q26575 methanobact
104	7	4.0	619	2	Q8RKJ2	Q8rkj2 bacillus su	177	6	3.4	86	2	Q9AJ72	Q9aj72 pseudomonas
105	7	4.0	629	11	Q8BTW0	Q8btw0 mus musculu	178	6	3.4	90	10	Q93WW8	Q93ww8 narcissus p
106	7	4.0	630	11	Q99K25	Q99k25 mus musculu	179	6	3.4	91	10	Q94KY3	Q94ky3 brassica ca
107	7	4.0	649	2	Q9R6F3	Q9r6f3 agrobacteri	180	6	3.4	92	16	Q7WHT8	Q7wht8 bordetella
108	7	4.0	650	5	Q86B43	Q86b43 drosophila	181	6	3.4	92	16	Q7W6V0	Q7w6v0 bordetella
109	7	4.0	669	10	Q9M9W6	Q9m9w6 arabidopsis	182	6	3.4	94	2	P97101	P97101 alcaligenes
110	7	4.0	669	10	Q8GX86	Q8gx86 arabidopsis	183	6	3.4	95	10	Q94KY6	Q94ky6 brassica na
111	7	4.0	669	16	Q8DKV3	Q8dkv3 synechococc	184	6	3.4	95	10	Q94KY4	Q94ky4 brassica ol
112	7	4.0	741	5	Q9VCU3	Q9vcu3 drosophila	185	6	3.4	96	10	Q9M3U0	Q9m3u0 betula verr
113	7	4.0	765	11	Q9R055	Q9r055 mus musculu	186	6	3.4	97	10	Q94KY8	Q94ky8 brassica na
114	7	4.0	782	5	Q61073	Q61073 paramecium	187	6	3.4	99	11	Q8BTB2	Q8btb2 mus musculu
115	7	4.0	783	12	Q91331	Q91331 cercopithic	188	6	3.4	100	16	Q981Y3	Q981y3 rhizobium l
116	7	4.0	838	5	Q9VQA9	Q9vqa9 drosophila	189	6	3.4	101	2	Q93EH9	Q93eh9 helicobacte
117	7	4.0	844	16	Q82RP4	Q82rp4 streptomyce	190	6	3.4	101	11	Q8CDW4	Q8cdw4 mus musculu
118	7	4.0	853	4	Q14333	Q14333 homo sapien	191	6	3.4	102	4	Q8WY95	Q8wy95 homo sapien
119	7	4.0	881	16	Q8XQC2	Q8xqc2 ralstonia s	192	6	3.4	102	4	Q9BW61	Q9bw61 homo sapien
120	7	4.0	889	5	Q9VZ63	Q9vz63 drosophila	193	6	3.4	102	11	Q9D9Z5	Q9d9z5 mus musculu
121	7	4.0	894	11	Q9ER61	Q9er61 mus musculu	194	6	3.4	102	11	Q8C870	Q8c870 mus musculu
122	7	4.0	894	11	Q80VU3	Q80vu3 mus musculu	195	6	3.4	102	16	Q92J02	Q92j02 rickettsia
123	7	4.0	894	11	Q8BJS8	Q8bjs8 mus musculu	196	6	3.4	102	16	Q88G24	Q88g24 pseudomonas
124	7	4.0	920	5	Q9GRN4	Q9grn4 leishmania	197	6	3.4	104	2	Q849V4	Q849v4 escherichia
125	7	4.0	948	16	Q9HZE4	Q9hze4 pseudomonas	198	6	3.4	104	2	Q49096	Q49096 mycoplasma
126	7	4.0	950	4	Q8NHY8	Q8nhy8 homo sapien	199	6	3.4	104	13	Q7T2A3	Q7t2a3 brachydanio
127	7	4.0	967	16	Q8PD22	Q8pd22 xanthomonas	200	6	3.4	104	17	Q9YBH3	Q9ybh3 aeropyrum p
128	7	4.0	1036	5	Q76974	Q76974 paramecium	201	6	3.4	104	17	Q972E9	Q972e9 sulfolobus
129	7	4.0	1037	5	Q9N9D8	Q9n9d8 paramecium	202	6	3.4	106	5	Q8WPQ7	Q8wpq7 trypanosoma
130	7	4.0	1047	10	Q8VZX6	Q8vzx6 medicago tr	203	6	3.4	107	16	Q8DKX9	Q8dkx9 synechococc
131	7	4.0	1048	10	Q42883	Q42883 lycopersico	204	6	3.4	108	2	Q8KUX3	Q8kux3 synechococc
132	7	4.0	1186	3	Q12466	Q12466 saccharomyc	205	6	3.4	111	3	P87273	P87273 saccharomyc
133	7	4.0	1217	10	Q9SRY2	Q9sry2 arabidopsis	206	6	3.4	112	16	Q97H90	Q97h90 clostridium
134	7	4.0	1282	4	Q8TER0	Q8ter0 homo sapien	207	6	3.4	112	16	Q97H90	Q97h90 clostridium
135	7	4.0	1516	3	Q9Y752	Q9y752 candida alb	208	6	3.4	113	10	Q8S1F9	Q8s1f9 oryza sativ
136	7	4.0	1595	2	Q52373	Q52373 caldicellul	209	6	3.4	114	10	Q84UN6	Q84un6 oryza sativ
137	7	4.0	1764	2	Q93T34	Q93t34 haemophilus	210	6	3.4	115	10	Q8S8K3	Q8s8k3 arabidopsis
138	7	4.0	1860	4	Q8IZC6	Q8izc6 homo sapien	211	6	3.4	116	2	Q93KX9	Q93kx9 streptomyce
139	7	4.0	2108	16	Q53901	Q53901 mycobacteri	212	6	3.4	116	16	Q7U7D8	Q7u7d8 synechococc
140	7	4.0	2108	16	Q7VEZ3	Q7vez3 mycobacteri	213	6	3.4	117	2	Q9KWH8	Q9kwh8 streptomyce
141	7	4.0	2616	5	Q9VRX5	Q9vrx5 drosophila	214	6	3.4	117	11	Q9CV32	Q9cv32 mus musculu
142	6	3.4	26	4	Q9UGK1	Q9ugk1 homo sapien	215	6	3.4	117	16	Q97G06	Q97g06 clostridium
143	6	3.4	35	4	Q9BVQ6	Q9bvq6 homo sapien	216	6	3.4	118	10	Q9SET5	Q9set5 witheringia
144	6	3.4	38	2	Q93I53	Q93i53 bacillus su	217	6	3.4	119	16	Q9X8X4	Q9x8x4 streptomyce
145	6	3.4	69	12	Q8QXG6	Q8qxg6 human coxsa	218	6	3.4	120	12	Q8S685	Q8s685 reovirus sp
146	6	3.4	69	12	Q8QWB8	Q8qwb8 human coxsa	219	6	3.4	120	12	Q85698	Q85698 reovirus sp
147	6	3.4	69	12	Q8QWB7	Q8qwb7 human coxsa	220	6	3.4	120	12	Q85696	Q85696 reovirus sp
148	6	3.4	69	12	Q8QXG5	Q8qxg5 human coxsa	221	6	3.4	121	2	Q9RMS0	Q9rms0 bradyrhizob
149	6	3.4	69	12	Q8QXG4	Q8qxg4 human coxsa	222	6	3.4	121	10	Q41308	Q41308 solanum car
150	6	3.4	69	12	Q7TB72	Q7tb72 human coxsa	223	6	3.4	121	16	Q8NPD8	Q8npd8 corynebacte
151	6	3.4	69	12	Q7TB71	Q7tb71 human coxsa	224	6	3.4	123	2	Q8KLE4	Q8kle4 rhizobium e
152	6	3.4	69	12	Q7TB70	Q7tb70 human coxsa	225	6	3.4	123	16	Q922F6	Q922f6 rhizobium m
153	6	3.4	69	12	Q7TB69	Q7tb69 human coxsa	226	6	3.4	123	16	Q8PEL1	Q8pell xanthomonas
154	6	3.4	69	12	Q7TB68	Q7tb68 human coxsa	227	6	3.4	124	2	Q7WYD8	Q7wxd8 alcaligenes
155	6	3.4	69	12	Q7TB67	Q7tb67 human coxsa	228	6	3.4	124	16	Q86818	Q86818 streptomyce
156	6	3.4	69	12	Q7TB66	Q7tb66 human coxsa	229	6	3.4	125	12	Q91T52	Q91t52 newcastle d
157	6	3.4	69	12	Q7TB65	Q7tb65 human coxsa	230	6	3.4	125	12	Q91T48	Q91t48 newcastle d
158	6	3.4	69	12	Q7TB64	Q7tb64 human coxsa	231	6	3.4	126	11	Q8K1D0	Q8kld0 mus musculu
159	6	3.4	69	12	Q7TB63	Q7tb63 human coxsa	232	6	3.4	127	5	Q86EK8	Q86ek8 schistosoma
160	6	3.4	69	12	Q7TB62	Q7tb62 human coxsa	233	6	3.4	127	10	Q8LG35	Q8lg35 arabidopsis
161	6	3.4	69	12	Q7TB61	Q7tb61 human coxsa	234	6	3.4	127	16	Q92I66	Q92i66 rickettsia
162	6	3.4	69	12	Q7TB60	Q7tb60 human coxsa	235	6	3.4	127	16	Q83PC4	Q83pc4 shigella fl
										127	16	Q7UB98	Q7ub98 shigella fl



236	6	3.4	128	2	Q8GBJ7	Q8gbj7 gamma-prote	309	6	3.4	172	16	Q9PCL6	Q9pcl6 xylella fas
237	6	3.4	129	5	Q86FG0	Q86fg0 schistosoma	310	6	3.4	172	16	Q9L1Y8	Q9lly8 streptomyce
238	6	3.4	129	16	P71839	P71839 mycobacteri	311	6	3.4	173	11	Q8BHZ8	Q8bhz8 mus musculu
239	6	3.4	129	16	Q7U1A1	Q7ula1 mycobacteri	312	6	3.4	173	12	Q80GW6	Q80gw6 hepatitis b
240	6	3.4	130	10	Q949F4	Q949f4 oryza sativ	313	6	3.4	173	16	Q9A9B1	Q9a9b1 caulobacter
241	6	3.4	130	11	Q8BL61	Q8bl61 mus musculu	314	6	3.4	173	16	Q9RI59	Q9ri59 streptomyce
242	6	3.4	130	16	Q7UEI2	Q7uei2 rhodopirell	315	6	3.4	173	16	Q8CNZ6	Q8cnz6 staphylococ
243	6	3.4	132	4	Q8NB79	Q8nb79 homo sapien	316	6	3.4	174	4	Q8N7F7	Q8n7f7 homo sapien
244	6	3.4	132	11	Q8BMZ9	Q8bmz9 mus musculu	317	6	3.4	174	6	Q8WNT5	Q8wnt5 canis lupus
245	6	3.4	133	5	Q9U241	Q9u241 caenorhabdi	318	6	3.4	175	16	Q9I2M8	Q9i2m8 pseudomonas
246	6	3.4	133	12	Q90704	Q90704 chicken ane	319	6	3.4	176	2	Q54195	Q54195 streptomyce
247	6	3.4	135	6	Q9N2E6	Q9n2e6 pongo pygma	320	6	3.4	176	10	Q7XNC7	Q7xnc7 oryza sativ
248	6	3.4	135	10	Q8LP83	Q8lp83 lycopersico	321	6	3.4	176	12	Q9EN83	Q9en83 newcasttle d
249	6	3.4	135	10	Q8LLM6	Q8llm6 triticum ae	322	6	3.4	176	16	Q9KD61	Q9kd61 bacillus ha
250	6	3.4	135	10	Q8LHF1	Q8lhf1 oryza sativ	323	6	3.4	176	16	Q8G6Q4	Q8g6q4 bifidobacte
251	6	3.4	135	10	Q8H328	Q8h328 oryza sativ	324	6	3.4	176	16	Q82CF7	Q82cf7 streptomyce
252	6	3.4	135	10	Q7XB38	Q7xb38 capsicum an	325	6	3.4	177	5	Q8STS3	Q8sts3 encephalito
253	6	3.4	135	16	Q9PGT9	Q9pgt9 xylella fas	326	6	3.4	177	5	Q9XWQ9	Q9xwq9 caenorhabdi
254	6	3.4	135	16	Q7UUT5	Q7uut5 rhodopirell	327	6	3.4	177	17	Q9YF03	Q9yfo3 aeropyrum p
255	6	3.4	138	12	Q9DWX7	Q9dwx7 human coxsa	328	6	3.4	177	5	Q9XVA3	Q9xva3 caenorhabdi
256	6	3.4	139	2	O83014	O83014 streptomyce	329	6	3.4	178	10	Q8RXV9	Q8rxv9 arabidopsi
257	6	3.4	139	14	Q99IU1	Q99iu1 uncultured	330	6	3.4	180	13	Q9IBA1	Q9iba1 potamotrygo
258	6	3.4	139	16	Q98F74	Q98f74 rhizobium l	331	6	3.4	180	16	Q8GLD9	Q8gld9 bruceella su
259	6	3.4	140	16	Q9CI72	Q9ci72 lactococcus	332	6	3.4	180	16	Q8DWK7	Q8dwk7 streptococc
260	6	3.4	141	2	O30613	O30613 myxococcus	333	6	3.4	180	16	Q89IX2	Q89ix2 bradyrhizob
261	6	3.4	141	16	Q46850	Q46850 escherichia	334	6	3.4	181	10	Q8LP95	Q8lp95 lycopersico
262	6	3.4	141	16	Q82QS0	Q82qs0 streptomyce	335	6	3.4	182	16	Q9CC07	Q9cc07 mycobacteri
263	6	3.4	142	10	Q93WC6	Q93wc6 arabidopsi	336	6	3.4	182	16	Q87N88	Q87n88 vibrio para
264	6	3.4	142	12	Q69341	Q69341 suid herpes	337	6	3.4	182	17	Q8TWT2	Q8twt2 methanopyru
265	6	3.4	144	2	Q9ZH81	Q9zh81 mycobacteri	338	6	3.4	183	10	Q9ASL4	Q9asl4 oryza sativ
266	6	3.4	144	10	Q7XZ59	Q7xz59 griffithsia	339	6	3.4	183	16	Q89MW1	Q89mz1 bradyrhizob
267	6	3.4	144	12	Q80GW7	Q80gw7 hepatitis b	340	6	3.4	184	16	Q8PCR1	Q8pcr1 xanthomonas
268	6	3.4	144	16	Q9RTZ1	Q9rtz1 deinococcus	341	6	3.4	184	16	Q7V8A1	Q7v8a1 prochloroco
269	6	3.4	145	16	Q8FUD2	Q8fud2 corynebacte	342	6	3.4	185	16	Q55888	Q55888 synchocyst
270	6	3.4	146	2	Q93P93	Q93p93 microscilla	343	6	3.4	186	5	Q9VS97	Q9vs97 drosophila
271	6	3.4	147	4	Q9H657	Q9h657 homo sapien	344	6	3.4	187	16	Q9JZK6	Q9jzk6 neisseria m
272	6	3.4	147	17	Q8TXN3	Q8txn3 methanopyru	345	6	3.4	187	16	Q9A694	Q9a694 caulobacter
273	6	3.4	148	16	Q92NQ2	Q92nq2 rhizobium m	346	6	3.4	188	2	Q9RQ72	Q9rq72 bacteroides
274	6	3.4	149	16	Q8A8N5	Q8a8n5 bacteroides	347	6	3.4	188	16	Q9RUC4	Q9ruc4 deinococcus
275	6	3.4	150	5	Q95WT0	Q95wt0 trypanosoma	348	6	3.4	189	2	Q83WD5	Q83wd5 streptomyce
276	6	3.4	150	12	Q910T6	Q910t6 lumpy skin	349	6	3.4	189	10	Q9FPF1	Q9fpf1 pinus strob
277	6	3.4	150	16	Q87PD6	Q87pd6 vibrio para	350	6	3.4	189	16	O53612	O53612 mycobacteri
278	6	3.4	150	16	Q81D46	Q81d46 bacillus ce	351	6	3.4	189	16	Q7U2X5	Q7u2x5 mycobacteri
279	6	3.4	151	5	Q8WQN6	Q8wqn6 anopheles g	352	6	3.4	190	5	Q9VV97	Q9vv97 drosophila
280	6	3.4	151	10	Q93XB2	Q93xb2 pisum sativ	353	6	3.4	190	16	Q9F391	Q9f391 streptomyce
281	6	3.4	152	2	Q8RLD9	Q8rl d9 mycoplasma	354	6	3.4	191	4	Q96C20	Q96c20 homo sapien
282	6	3.4	152	2	Q9F2A3	Q9f2a3 chromatiu	355	6	3.4	192	15	O56014	O56014 human immun
283	6	3.4	152	2	Q50820	Q50820 mycobacteri	356	6	3.4	192	15	Q9IQD5	Q9iqd5 human immun
284	6	3.4	153	4	Q8NHH4	Q8nhh4 homo sapien	357	6	3.4	192	15	Q9IQD2	Q9iqd2 human immun
285	6	3.4	153	16	Q9S1Q7	Q9s1q7 streptomyce	358	6	3.4	192	15	O55997	O55997 human immun
286	6	3.4	155	16	Q8XLQ9	Q8xlg9 clostridium	359	6	3.4	192	15	Q9Q5K7	Q9q5k7 human immun
287	6	3.4	156	2	Q8VQD5	Q8vqd5 thermus the	360	6	3.4	192	15	O56008	O56008 human immun
288	6	3.4	156	12	Q8B4K2	Q8b4k2 human parec	361	6	3.4	192	15	O55999	O55999 human immun
289	6	3.4	157	2	Q8VQD7	Q8vqd7 thermus aqu	362	6	3.4	192	15	O56002	O56002 human immun
290	6	3.4	157	3	Q06051	Q06051 saccharomyc	363	6	3.4	192	15	Q9IQD4	Q9iqd4 human immun
291	6	3.4	157	4	Q8N7Z9	Q8n7z9 homo sapien	364	6	3.4	192	15	O55998	O55998 human immun
292	6	3.4	160	16	Q8UEE7	Q8uee7 agrobacteri	365	6	3.4	192	15	O56000	O56000 human immun
293	6	3.4	160	17	Q8TJF7	Q8tjf7 methanosarc	366	6	3.4	192	15	O56013	O56013 human immun
294	6	3.4	161	5	Q9N9S6	Q9n9s6 leishmania	367	6	3.4	192	15	O56015	O56015 human immun
295	6	3.4	161	16	Q8DCH6	Q8dch6 vibrio vuln	368	6	3.4	192	15	O56007	O56007 human immun
296	6	3.4	161	16	Q82K38	Q82k38 streptomyce	369	6	3.4	192	15	Q8J9C3	Q8j9c3 human immun
297	6	3.4	162	4	Q9UK00	Q9uk00 homo sapien	370	6	3.4	192	15	O11951	O11951 human immun
298	6	3.4	162	16	Q9KUV3	Q9kuv3 vibrio chol	371	6	3.4	192	15	Q9YVF3	Q9yvf3 human immun
299	6	3.4	163	10	Q9SL93	Q9sl93 arabidopsi	372	6	3.4	192	15	O56009	O56009 human immun
300	6	3.4	165	2	Q56651	Q56651 vibrio chol	373	6	3.4	192	15	O71257	O71257 human immun
301	6	3.4	165	5	Q952K9	Q952k9 caenorhabdi	374	6	3.4	192	15	O56012	O56012 human immun
302	6	3.4	165	16	Q9X9W2	Q9x9w2 streptomyce	375	6	3.4	192	15	O56011	O56011 human immun
303	6	3.4	166	2	Q9AIT8	Q9ait8 ralstonia s	376	6	3.4	192	15	O56005	O56005 human immun
304	6	3.4	166	10	Q8S3X8	Q8s3x8 oryza sativ	377	6	3.4	192	15	Q73415	Q73415 human immun
305	6	3.4	166	17	Q8U0W5	Q8u0w5 pyrococcus	378	6	3.4	192	15	Q9IQD3	Q9iqd3 human immun
306	6	3.4	169	2	Q51528	Q51528 pseudomonas	379	6	3.4	192	15	Q73414	Q73414 human immun
307	6	3.4	169	16	Q9HVN0	Q9hvn0 pseudomonas	380	6	3.4	192	15	Q73417	Q73417 human immun
308	6	3.4	172	10	Q9FSL4	Q9fsl4 oryza sativ	381	6	3.4	192	15	Q73416	Q73416 human immun

Q9pcl6	xylella fas
Q9lly8	streptomyce
Q8bhz8	mus musculu
Q80gw6	hepatitis b
Q9a9b1	caulobacter
Q9ri59	streptomyce
Q8cnz6	staphylococ
Q8n7f7	homo sapien
Q8wnt5	canis lupus
Q9i2m8	pseudomonas
Q54195	streptomyce
Q7xnc7	oryza sativ
Q9en83	newcasttle d
Q9kd61	bacillus ha
Q8g6q4	bifidobacte
Q82cf7	streptomyce
Q8sts3	encephalito
Q9xwq9	caenorhabdi
Q9yfo3	aeropyrum p
Q9xva3	caenorhabdi
Q8rxv9	arabidopeis
Q9iba1	potamotrygo
Q8gld9	bruceella su
Q8dwk7	streptococc
Q89ix2	bradyrhizob
Q8lp95	lycopersico
Q9cc07	mycobacteri
Q87n88	vibrio para
Q8twt2	methanopyru
Q9asl4	oryza sativ
Q89mz1	bradyrhizob
Q8pcr1	xanthomonas
Q7v8a1	prochloroco
Q55888	synchocyst
Q9vs97	drosophila
Q9jzk6	neisseria m
Q9a694	caulobacter
Q9rq72	bacteroides
Q9ruc4	deinococcus
Q83wd5	streptomyce
Q9fpf1	pinus strob
O53612	mycobacteri
Q7u2x5	mycobacteri
Q9vv97	drosophila
Q9f391	streptomyce
Q96c20	homo sapien
O56014	human immun
Q9iqd5	human immun
Q9iqd2	human immun
O55997	human immun
Q9q5k7	human immun
O56008	human immun
O55999	human immun
O56002	human immun
Q9iqd4	human immun
O55998	human immun
O56000	human immun
O56013	human immun
O56015	human immun
O56007	human immun
Q8j9c3	human immun
O11951	human immun
Q9yvf3	human immun
O56009	human immun
O71257	human immun
O56012	human immun
O56011	human immun
O56005	human immun
Q73415	human immun
Q9iqd3	human immun
Q73414	human immun
Q73417	human immun
Q73416	human immun

382	6	3.4	192	15	056001	O56001 human immun	455	6	3.4	210	16	Q982Z6	Q982z6 rhizobium l
383	6	3.4	192	15	055946	O55946 human immun	456	6	3.4	210	16	Q9RCV9	Q9rcv9 streptomyce
384	6	3.4	192	15	056004	O56004 human immun	457	6	3.4	211	10	Q9M1F8	Q9mlf8 arabadopsis
385	6	3.4	192	15	055940	O55940 human immun	458	6	3.4	211	16	Q8ZMP9	Q8zmp9 salmonella
386	6	3.4	192	15	056003	O56003 human immun	459	6	3.4	211	16	Q8ZK78	Q8zk78 salmonella
387	6	3.4	192	15	056010	O56010 human immun	460	6	3.4	211	16	Q8Z161	Q8z161 salmonella
388	6	3.4	192	15	056006	O56006 human immun	461	6	3.4	212	16	Q92B55	Q92b55 listeria in
389	6	3.4	192	15	Q7SPQ1	Q7spq1 human immun	462	6	3.4	213	2	Q9RHU6	Q9rhu6 streptomyce
390	6	3.4	192	16	Q7UTS5	Q7uts5 rhodopirell	463	6	3.4	213	12	Q8JMH2	Q8jmh2 mamestra co
391	6	3.4	193	12	Q91GQ3	Q91gq3 kaposi's sa	464	6	3.4	213	16	Q8UC12	Q8uc12 agrobacteri
392	6	3.4	193	12	Q91GQ1	Q91gq1 kaposi's sa	465	6	3.4	213	16	Q9Z677	Q9z677 streptomyce
393	6	3.4	193	12	Q91GQ0	Q91gq0 kaposi's sa	466	6	3.4	213	16	Q82JH8	Q82jh8 streptomyce
394	6	3.4	193	12	Q91GQ2	Q91gq2 kaposi's sa	467	6	3.4	214	10	Q8LP80	Q8lp80 lycopersico
395	6	3.4	193	12	Q91GP8	Q91gp8 kaposi's sa	468	6	3.4	214	10	Q8LP85	Q8lp85 lycopersico
396	6	3.4	193	12	Q91GP9	Q91gp9 kaposi's sa	469	6	3.4	215	5	Q9VGN3	Q9vgn3 drosophila
397	6	3.4	193	12	Q91GQ4	Q91gq4 kaposi's sa	470	6	3.4	215	5	Q86JI4	Q86ji4 dictyosteli
398	6	3.4	193	12	Q91GP7	Q91gp7 kaposi's sa	471	6	3.4	215	16	Q9RZR6	Q9rrzr6 deinococcus
399	6	3.4	193	12	Q91GP6	Q91gp6 kaposi's sa	472	6	3.4	215	16	Q8PM24	Q8pm24 xanthomonas
400	6	3.4	193	16	Q8YGI1	Q8ygi1 brucella me	473	6	3.4	216	4	O43871	O43871 homo sapien
401	6	3.4	193	16	Q8P3J6	Q8p3j6 xanthomonas	474	6	3.4	216	4	Q9BU55	Q9bu55 homo sapien
402	6	3.4	193	16	Q89KS5	Q89ks5 bradyrhizob	475	6	3.4	216	10	Q42441	Q42441 solanum cha
403	6	3.4	194	2	Q93F30	Q93f30 shigella fl	476	6	3.4	216	12	Q8QLG3	Q8qlg3 mamestra co
404	6	3.4	194	4	Q9HC43	Q9hc43 homo sapien	477	6	3.4	217	2	Q9RIM0	Q9rim0 streptomyce
405	6	3.4	194	16	Q83NR4	Q83nr4 tropheryma	478	6	3.4	217	5	Q86JI2	Q86ji2 dictyosteli
406	6	3.4	194	16	Q83FW8	Q83fw8 tropheryma	479	6	3.4	217	11	Q8BX71	Q8bx71 mus musculu
407	6	3.4	195	5	Q8WQY7	Q8wqy7 aequorea ma	480	6	3.4	218	2	O54227	O54227 streptomyce
408	6	3.4	196	4	Q9NSL9	Q9nsl9 homo sapien	481	6	3.4	218	12	Q7TFH4	Q7tfh4 rhesus cyto
409	6	3.4	196	5	Q7YTX0	Q7ytx0 aequorea vi	482	6	3.4	218	16	Q8XVC6	Q8xvc6 ralstonia s
410	6	3.4	196	5	Q7YTW9	Q7ytw9 aequorea vi	483	6	3.4	218	16	Q87C38	Q87c38 xylella fas
411	6	3.4	196	5	Q7YTW8	Q7ytw8 aequorea vi	484	6	3.4	218	16	Q7UJW3	Q7ujw3 rhodopirell
412	6	3.4	196	5	Q7YTW7	Q7ytw7 aequorea vi	485	6	3.4	219	11	Q8C3Z7	Q8c3z7 mus musculu
413	6	3.4	196	12	Q7THE8	Q7the8 bovine kobu	486	6	3.4	219	16	Q9RZG4	Q9rzg4 deinococcus
414	6	3.4	196	12	Q7THE7	Q7the7 bovine kobu	487	6	3.4	220	11	Q8C3Y1	Q8c3y1 mus musculu
415	6	3.4	196	12	Q7THE6	Q7the6 bovine kobu	488	6	3.4	220	11	Q80VX8	Q80vx8 mus musculu
416	6	3.4	196	12	Q7THE5	Q7the5 bovine kobu	489	6	3.4	220	16	Q9A1I6	Q9ali6 streptococc
417	6	3.4	196	12	Q7THE4	Q7the4 bovine kobu	490	6	3.4	220	16	Q8P420	Q8p420 xanthomonas
418	6	3.4	196	12	Q7THE3	Q7the3 bovine kobu	491	6	3.4	220	16	Q8P2P4	Q8p2p4 streptococc
419	6	3.4	196	12	Q7THE2	Q7the2 bovine kobu	492	6	3.4	220	16	Q9RUD2	Q9rjd2 streptomyce
420	6	3.4	196	12	Q7THE1	Q7the1 bovine kobu	493	6	3.4	220	16	Q8X8P1	Q8k8p1 streptococc
421	6	3.4	196	12	Q7THE0	Q7the0 bovine kobu	494	6	3.4	220	16	Q8A7F8	Q8a7f8 bacteroides
422	6	3.4	196	12	Q7THD9	Q7thd9 bovine kobu	495	6	3.4	221	10	Q8S9S2	Q8s9s2 oryza sativ
423	6	3.4	196	12	Q7THD8	Q7thd8 bovine kobu	496	6	3.4	222	10	P92970	P92970 arabadopsis
424	6	3.4	196	12	Q7THD7	Q7thd7 bovine kobu	497	6	3.4	222	11	Q8BNT6	Q8bnt6 mus musculu
425	6	3.4	197	16	Q8YFG1	Q8yfg1 brucella me	498	6	3.4	222	16	Q7U5R9	Q7u5r9 synechococc
426	6	3.4	197	16	Q7UUP1	Q7uup1 rhodopirell	499	6	3.4	223	10	Q94JB9	Q94jb9 oryza sativ
427	6	3.4	199	2	Q9RCT0	Q9rct0 synechococc	500	6	3.4	223	16	Q9KXW5	Q9kxw5 streptomyce
428	6	3.4	199	5	Q95T09	Q95t09 drosophila	501	6	3.4	224	5	Q86AE9	Q86ae9 dictyosteli
429	6	3.4	199	16	Q9RZL3	Q9rzl3 deinococcus	502	6	3.4	224	16	Q9L2I3	Q9l2i3 streptomyce
430	6	3.4	199	16	Q98BP9	Q98bp9 rhizobium l	503	6	3.4	224	16	Q88B84	Q88b84 pseudomonas
431	6	3.4	199	16	Q8ZLN10	Q8zn10 salmonella	504	6	3.4	224	16	Q7WIP8	Q7wip8 bordetella
432	6	3.4	200	10	Q9LZP4	Q9lzp4 arabadopsis	505	6	3.4	224	16	Q7W736	Q7w736 bordetella
433	6	3.4	201	5	Q19830	Q19830 caenorhabdi	506	6	3.4	225	8	Q7YAZ3	Q7yaz3 cordylosaur
434	6	3.4	201	10	Q9M1F7	Q9mlf7 arabadopsis	507	6	3.4	225	10	Q949B2	Q949b2 oryza sativ
435	6	3.4	201	16	Q8ZPM7	Q8zpm7 salmonella	508	6	3.4	227	5	Q8T1I3	Q8t1i3 bombyx mori
436	6	3.4	201	16	Q8Z6Q4	Q8z6q4 salmonella	509	6	3.4	227	16	Q7WCW3	Q7wcw3 bordetella
437	6	3.4	202	5	Q9BLE7	Q9ble7 entamoeba h	510	6	3.4	227	16	Q7W5C9	Q7w5c9 bordetella
438	6	3.4	202	16	Q8ZPH2	Q8zph2 salmonella	511	6	3.4	227	16	Q7VZN2	Q7vzn2 bordetella
439	6	3.4	202	16	O06330	O06330 mycobacteri	512	6	3.4	228	16	Q89NM2	Q89nm2 bradyrhizob
440	6	3.4	202	16	Q8Z711	Q8z711 salmonella	513	6	3.4	229	16	Q8YET1	Q8yet1 brucella me
441	6	3.4	202	16	Q81D79	Q81d79 bacillus ce	514	6	3.4	230	2	Q8GP93	Q8gp93 streptococc
442	6	3.4	202	16	Q7TWG6	Q7twg6 mycobacteri	515	6	3.4	230	3	Q874B1	Q874b1 cordyceps m
443	6	3.4	203	2	O33170	O33170 mycobacteri	516	6	3.4	230	4	Q8WYK4	Q8wyk4 homo sapien
444	6	3.4	203	16	Q8UG60	Q8ug60 agrobacteri	517	6	3.4	230	4	Q8N8W8	Q8nbw8 homo sapien
445	6	3.4	203	16	Q87D87	Q87d87 xylella fas	518	6	3.4	230	10	Q8H8F0	Q8h8f0 oryza sativ
446	6	3.4	204	16	Q9A298	Q9a298 caulobacter	519	6	3.4	230	16	Q8VEP8	Q8yep8 brucella me
447	6	3.4	204	16	Q8G2F6	Q8g2f6 brucella su	520	6	3.4	230	16	Q8G334	Q8g334 brucella su
448	6	3.4	205	4	Q96EI7	Q96ei7 homo sapien	521	6	3.4	231	16	Q8QB7	Q88qb7 pseudomonas
449	6	3.4	205	10	Q9ASL2	Q9asl2 oryza sativ	522	6	3.4	232	2	O06518	O06518 escherichia
450	6	3.4	205	11	Q7TMK5	Q7tmk5 mus musculu	523	6	3.4	232	5	Q8INK6	Q8ink6 drosophila
451	6	3.4	207	16	Q9ACR4	Q9acr4 streptomyce	524	6	3.4	232	16	Q98P48	Q98p48 rhizobium l
452	6	3.4	208	4	Q96KA8	Q96ka8 homo sapien	525	6	3.4	232	16	Q82JK5	Q82jk5 streptomyce
453	6	3.4	209	5	Q8SVJ4	Q8svj4 encephalito	526	6	3.4	233	2	Q8KJ84	Q8kj84 rhizobium l
454	6	3.4	210	10	Q7YOR7	Q7yor7 lycopersico	527	6	3.4	233	2	Q8GI84	Q8gi84 streptomyce



528	6	3.4	233	16	Q8G2J9	Q8g2j9 brucella su
529	6	3.4	234	5	Q86F97	Q86f97 schistosoma
530	6	3.4	234	17	Q96XK6	Q96xk6 sulfolobus
531	6	3.4	234	17	Q26747	Q26747 methanobact
532	6	3.4	235	2	Q9XCI4	Q9xcI4 shigella fl
533	6	3.4	235	10	Q04176	Q04176 brassica ca
534	6	3.4	235	16	Q9A5N9	Q9a5n9 caulobacter
535	6	3.4	235	16	Q07197	Q07197 mycobacteri
536	6	3.4	235	16	Q8PKD1	Q8pkd1 xanthomonas
537	6	3.4	235	16	Q88MA6	Q88ma6 pseudomonas
538	6	3.4	235	16	Q9Z7T6	Q9z7t6 chlamydia p
539	6	3.4	235	16	Q7TY32	Q7ty32 mycobacteri
540	6	3.4	236	5	Q86AU3	Q86au3 dictyosteli
541	6	3.4	236	16	Q88II7	Q88Ii7 pseudomonas
542	6	3.4	237	4	Q8NAK0	Q8nak0 homo sapien
543	6	3.4	237	5	Q8MXI7	Q8mxi7 caenorhabdi
544	6	3.4	237	10	Q04177	Q04177 brassica ca
545	6	3.4	237	16	Q8YFB9	Q8yfb9 brucella me
546	6	3.4	237	17	Q970M0	Q970m0 sulfolobus
547	6	3.4	238	9	Q38301	Q38301 lactococcus
548	6	3.4	238	10	Q8LPE5	Q8lpe5 cicer ariet
549	6	3.4	238	16	Q8PNN3	Q8pnn3 xanthomonas
550	6	3.4	238	16	Q8FSW8	Q8fsw8 corynebacte
551	6	3.4	238	17	Q28437	Q28437 archaeoglob
552	6	3.4	239	16	Q8UK52	Q8uk52 agrobacteri
553	6	3.4	239	16	Q88GB7	Q88gb7 pseudomonas
554	6	3.4	240	4	Q96DY3	Q96dy3 homo sapien
555	6	3.4	240	16	Q8DMQ8	Q8dmq8 synechococc
556	6	3.4	241	10	Q9FJD8	Q9fjd8 arabidopsis
557	6	3.4	241	16	Q8XV65	Q8xv65 ralstonia s
558	6	3.4	241	16	Q828D2	Q828d2 streptomyce
559	6	3.4	242	10	Q7Y1X0	Q7y1x0 porphyra ye
560	6	3.4	242	16	Q8RGQ1	Q8rgq1 fusobacteri
561	6	3.4	242	16	Q81RT0	Q81rt0 bacillus an
562	6	3.4	242	16	Q81EL1	Q81el1 bacillus ce
563	6	3.4	242	16	Q7U774	Q7u774 synechococc
564	6	3.4	245	10	Q81603	Q81603 mesembryant
565	6	3.4	247	10	Q8GT26	Q8gt26 galdieria p
566	6	3.4	247	10	Q9ATD9	Q9atd9 gossypium h
567	6	3.4	248	16	Q8FNU4	Q8fnu4 corynebacte
568	6	3.4	249	5	Q44399	Q44399 caenorhabdi
569	6	3.4	249	10	Q38850	Q38850 arabidopsis
570	6	3.4	249	11	Q8C1Q8	Q8clq8 mus musculu
571	6	3.4	249	16	Q98IY4	Q98iy4 rhizobium l
572	6	3.4	249	16	Q7WPC6	Q7wpc6 bordetella
573	6	3.4	249	16	Q7VUB4	Q7vub4 bordetella
574	6	3.4	249	17	Q9HNY3	Q9hny3 halobacteri
575	6	3.4	249	17	Q8ZT00	Q8zt00 pyrobaculum
576	6	3.4	250	2	Q7WX68	Q7wx68 alcaligenes
577	6	3.4	250	10	Q43824	Q43824 raphanus sa
578	6	3.4	250	10	Q49122	Q49122 brassica ju
579	6	3.4	250	10	Q65161	Q65161 zantedeschi
580	6	3.4	250	10	Q8W3J6	Q8w3j6 brassica ol
581	6	3.4	250	10	Q8W3J5	Q8w3j5 brássica ol
582	6	3.4	250	10	Q9SMD3	Q9smd3 lycopersico
583	6	3.4	250	10	P93657	P93657 brassica na
584	6	3.4	250	10	Q8H9F0	Q8h9f0 solanum tub
585	6	3.4	250	10	Q8H6F4	Q8h6f4 brassica ju
586	6	3.4	250	10	Q8H6F3	Q8h6f3 brassica ju
587	6	3.4	250	10	Q8GU36	Q8gu36 physcomitre
588	6	3.4	250	10	Q7XAT8	Q7xat8 crocus sati
589	6	3.4	250	10	Q7X6B7	Q7x6b7 oryza sativ
590	6	3.4	250	16	Q89SB3	Q89sb3 bradyrhizob
591	6	3.4	250	16	Q892R3	Q892r3 clostridium
592	6	3.4	250	16	Q88CM0	Q88cm0 pseudomonas
593	6	3.4	250	16	Q83M04	Q83m04 shigella fl
594	6	3.4	251	8	P92543	P92543 arabidopsis
595	6	3.4	252	5	Q86P09	Q86p09 drosophila
596	6	3.4	252	17	Q9YC93	Q9yc93 aeropyrum p
597	6	3.4	252	17	Q9HRZ9	Q9hrz9 halobacteri
598	6	3.4	253	5	Q8T940	Q8t940 junonia coe
599	6	3.4	253	16	Q8U759	Q8u759 agrobacteri
600	6	3.4	253	16	Q9KYU7	Q9kyu7 streptomyce

601	6	3.4	253	16	Q89TM5	Q89tm5 bradyrhizob
602	6	3.4	253	16	Q82H56	Q82h56 streptomyce
603	6	3.4	254	10	Q81604	Q81604 mesembryant
604	6	3.4	254	16	Q8PL43	Q8pl43 xanthomonas
605	6	3.4	254	16	Q8DJS3	Q8djs3 synechococc
606	6	3.4	255	4	Q8N9S2	Q8n9s2 homo sapien
607	6	3.4	255	10	Q8LR75	Q8lr75 oryza sativ
608	6	3.4	255	11	Q8C9A2	Q8c9a2 mus musculu
609	6	3.4	255	12	Q89459	Q89459 heliothis a
610	6	3.4	255	16	Q9WYS2	Q9wys2 thermotoga
611	6	3.4	255	17	Q9Y9V5	Q9y9y5 aeropyrum p
612	6	3.4	256	16	Q8XR32	Q8xr32 ralstonia s
613	6	3.4	256	16	Q89FP8	Q89fp8 bradyrhizob
614	6	3.4	257	13	Q7ZVE3	Q7zve3 brachydanio
615	6	3.4	257	16	Q9HV39	Q9hvv39 pseudomonas
616	6	3.4	257	16	Q930L3	Q930l3 rhizobium m
617	6	3.4	258	2	Q8KTI5	Q8kti5 azotobacter
618	6	3.4	258	10	Q8LQ00	Q8lqq0 oryza sativ
619	6	3.4	258	13	Q90Y62	Q90y62 brachydanio
620	6	3.4	258	17	Q9HHR7	Q9hhr7 halobacteri
621	6	3.4	259	10	Q93XK8	Q93xk8 lycopersico
622	6	3.4	259	16	Q9K7D6	Q9k7d6 bacillus ha
623	6	3.4	260	2	Q8GDY3	Q8gdy3 heliobacill
624	6	3.4	260	5	Q8SSU2	Q8ssu2 dictyosteli
625	6	3.4	260	10	Q42909	Q42909 mesembryant
626	6	3.4	260	16	Q8NTR2	Q8nt23 corynebacte
627	6	3.4	261	2	Q9ANB1	Q9anb1 bradyrhizob
628	6	3.4	262	5	Q9BL19	Q9bll9 bombyx mand
629	6	3.4	262	11	Q8R0X6	Q8r0x6 mus musculu
630	6	3.4	262	16	Q9X197	Q9xl97 thermotoga
631	6	3.4	263	4	Q8TCM0	Q8tcm0 homo sapien
632	6	3.4	263	16	Q9K3U8	Q9k3u8 streptomyce
633	6	3.4	263	16	Q8FWS7	Q8fws7 brucella su
634	6	3.4	263	16	Q82GP5	Q82gp5 streptomyce
635	6	3.4	263	17	Q9HHM2	Q9hhm2 halobacteri
636	6	3.4	264	11	Q80VX7	Q80vx7 mus musculu
637	6	3.4	265	3	P87092	P87092 emericella
638	6	3.4	265	2	Q77581	Q77581 bos taurus
639	6	3.4	266	2	Q56187	Q56187 streptomyce
640	6	3.4	266	16	Q8P7I0	Q8p7i0 xanthomonas
641	6	3.4	266	16	Q7WP74	Q7wp74 bordetella
642	6	3.4	266	16	Q7W1G4	Q7wl1g4 bordetella
643	6	3.4	266	17	Q971X8	Q971x8 sulfolobus
644	6	3.4	267	2	Q84HA9	Q84ha9 streptomyce
645	6	3.4	268	4	Q9NSS9	Q9nss9 homo sapien
646	6	3.4	268	12	Q9IC69	Q9ic69 kaposi's sa
647	6	3.4	268	12	Q9IC56	Q9ic56 kaposi's sa
648	6	3.4	268	12	Q9IC59	Q9ic59 kaposi's sa
649	6	3.4	268	12	Q9IC65	Q9ic65 kaposi's sa
650	6	3.4	268	12	Q9IC77	Q9ic77 kaposi's sa
651	6	3.4	268	12	Q9IC55	Q9ic55 kaposi's sa
652	6	3.4	268	12	Q9IC70	Q9ic70 kaposi's sa
653	6	3.4	268	12	Q9IC75	Q9ic75 kaposi's sa
654	6	3.4	268	12	Q9IC44	Q9ic44 kaposi's sa
655	6	3.4	268	12	Q9IC79	Q9ic79 kaposi's sa
656	6	3.4	268	12	Q9IC49	Q9ic49 kaposi's sa
657	6	3.4	268	12	Q9IC42	Q9ic42 kaposi's sa
658	6	3.4	268	12	Q9IC46	Q9ic46 kaposi's sa
659	6	3.4	268	12	Q9IC72	Q9ic72 kaposi's sa
660	6	3.4	268	12	Q9IC57	Q9ic57 kaposi's sa
661	6	3.4	268	12	Q9IC80	Q9ic80 kaposi's sa
662	6	3.4	268	12	Q9IC78	Q9ic78 kaposi's sa
663	6	3.4	268	12	Q9IC47	Q9ic47 kaposi's sa
664	6	3.4	268	12	Q9IC41	Q9ic41 kaposi's sa
665	6	3.4	268	12	Q9IC66	Q9ic66 kaposi's sa
666	6	3.4	268	12	Q9IC62	Q9ic62 kaposi's sa
667	6	3.4	268	12	Q9IC76	Q9ic76 kaposi's sa
668	6	3.4	268	12	Q9IC48	Q9ic48 kaposi's sa
669	6	3.4	268	12	Q9IC61	Q9ic61 kaposi's sa
670	6	3.4	268	12	Q9IC71	Q9ic71 kaposi's sa
671	6	3.4	268	12	Q9IC51	Q9ic51 kaposi's sa
672	6	3.4	268	12	Q9IC81	Q9ic81 kaposi's sa
673	6	3.4	268	12	Q9IC67	Q9ic67 kaposi's sa

674	6	3.4	268	12	Q9IC54	Q9ic54 kaposi's sa	747	6	3.4	288	11	Q8BR15	Q8br15 mus musculu
675	6	3.4	268	12	Q9IC64	Q9ic64 kaposi's sa	748	6	3.4	288	16	Q8YGG8	Q8ygg8 brucella me
676	6	3.4	268	12	Q9IC68	Q9ic68 kaposi's sa	749	6	3.4	288	16	Q8P6U9	Q8p6u9 xanthomonas
677	6	3.4	268	12	Q9IC45	Q9ic45 kaposi's sa	750	6	3.4	289	10	Q9ARD4	Q9ard4 hordeum vul
678	6	3.4	268	12	Q9IC73	Q9ic73 kaposi's sa	751	6	3.4	289	16	Q9A6U0	Q9a6u0 caulobacter
679	6	3.4	268	12	Q9IC60	Q9ic60 kaposi's sa	752	6	3.4	290	5	O45469	O45469 caenorhabdi
680	6	3.4	268	12	Q9IC58	Q9ic58 kaposi's sa	753	6	3.4	290	10	Q9SXB1	Q9sxb1 hordeum vul
681	6	3.4	268	12	Q9IC43	Q9ic43 kaposi's sa	754	6	3.4	291	4	Q9Y3A1	Q9y3a1 homo sapien
682	6	3.4	268	12	Q9IC63	Q9ic63 kaposi's sa	755	6	3.4	291	16	Q9JZL4	Q9jzl4 neisseria m
683	6	3.4	268	12	Q9IC50	Q9ic50 kaposi's sa	756	6	3.4	291	16	Q9JUN3	Q9jun3 neisseria m
684	6	3.4	268	12	Q9IC53	Q9ic53 kaposi's sa	757	6	3.4	291	16	Q8XYH0	Q8xyh0 ralstonia s
685	6	3.4	268	12	Q9IC74	Q9ic74 kaposi's sa	758	6	3.4	292	16	Q9AK27	Q9ak27 streptomyce
686	6	3.4	268	12	Q9IC52	Q9ic52 kaposi's sa	759	6	3.4	292	16	Q7W4D5	Q7wfu3 bordetella
687	6	3.4	268	16	Q8XS75	Q8xr36 corynebacte	760	6	3.4	292	16	Q7W4D5	Q7w4d5 bordetella
688	6	3.4	268	16	Q8NR36	Q8nr36 corynebacte	761	6	3.4	292	16	Q7W4D5	Q7w4d5 bordetella
689	6	3.4	269	12	Q9YMK9	Q9ymk9 lymantria d	762	6	3.4	293	4	O75391	O75391 homo sapien
690	6	3.4	269	16	Q92U77	Q92u77 rhizobium m	763	6	3.4	293	10	Q8HIP5	Q8hip5 glycine max
691	6	3.4	269	16	Q8FQ09	Q8fq09 corynebacte	764	6	3.4	294	2	Q93SI6	Q93si6 corynebacte
692	6	3.4	269	16	Q89MN5	Q89mn5 bradyrhizob	765	6	3.4	294	2	Q9RMF3	Q9rmf3 zymomonas m
693	6	3.4	270	16	Q8YM96	Q8ym96 anabaena sp	766	6	3.4	294	5	Q9VYK8	Q9vyk8 drosophila
694	6	3.4	271	2	Q9LVB5	Q9lbv5 streptomyce	767	6	3.4	294	10	Q7Y0S0	Q7y0s0 arabidopsis
695	6	3.4	271	5	Q9VPN7	Q9vpn7 drosophila	768	6	3.4	295	10	Q8LNY5	Q8lny5 nicotiana t
696	6	3.4	272	16	Q9HUU6	Q9hu6 pseudomonas	769	6	3.4	295	12	Q8V7I5	Q8v7i5 tt virus. o
697	6	3.4	272	16	Q8UDY8	Q8udy8 agrobacteri	770	6	3.4	296	2	Q9RN50	Q9rn50 mycobacteri
698	6	3.4	273	5	Q9U9P7	Q9u9p7 drosophila	771	6	3.4	296	2	Q9RN49	Q9rn49 mycobacteri
699	6	3.4	273	16	Q89MZ4	Q89mz4 bradyrhizob	772	6	3.4	296	2	Q9X657	Q9x657 streptomyce
700	6	3.4	273	16	Q88GG4	Q88gg4 pseudomonas	773	6	3.4	296	10	Q851Q3	Q851q3 oryza sativ
701	6	3.4	273	17	Q8ZSP1	Q8zsp1 pyrobaculum	774	6	3.4	297	10	Q43684	Q43684 vigna ungui
702	6	3.4	274	16	Q7WLS1	Q7wls1 bordetella	775	6	3.4	297	10	Q7YOR8	Q7yor8 lycopersico
703	6	3.4	275	1	Q8X240	Q8x240 haloarcula	776	6	3.4	297	17	O26454	O26454 methanobact
704	6	3.4	275	10	O23892	O23892 oryza sativ	777	6	3.4	297	17	Q9HSE7	Q9hse7 halobacteri
705	6	3.4	275	16	Q92FT0	Q92ft0 listeria in	778	6	3.4	298	2	Q8RS30	Q8rs30 comamonas a
706	6	3.4	275	16	Q8YAT9	Q8yat9 listeria mo	779	6	3.4	298	4	Q9NZE1	Q9nzel homo sapien
707	6	3.4	275	16	Q8NS72	Q8ns72 corynebacte	780	6	3.4	298	10	Q9S834	Q9s834 arabidopsis
708	6	3.4	275	17	Q9YCN0	Q9ycn0 aeropyrum p	781	6	3.4	298	11	Q8R587	Q8r587 mus musculu
709	6	3.4	276	5	Q17221	Q17221 bombyx mori	782	6	3.4	298	11	Q8K015	Q8k015 mus musculu
710	6	3.4	276	5	Q17222	Q17222 bombyx mori	783	6	3.4	298	11	Q80W01	Q80w01 mus musculu
711	6	3.4	276	10	Q8H274	Q8h274 oryza sativ	784	6	3.4	298	16	Q98KG0	Q98kg0 rhizobium l
712	6	3.4	276	16	Q9RYH3	Q9ryh3 deinococcus	785	6	3.4	298	16	Q8XU59	Q8xu59 ralstonia s
713	6	3.4	277	5	Q9NK49	Q9nk49 drosophila	786	6	3.4	298	16	Q9RD12	Q9rd12 streptomyce
714	6	3.4	277	10	Q9SB44	Q9sb44 arabidopsis	787	6	3.4	298	16	Q8ZH21	Q8zh21 yersinia pe
715	6	3.4	277	10	Q7X6W1	Q7x6w1 oryza sativ	788	6	3.4	299	5	Q9N871	Q9n871 leishmania
716	6	3.4	277	16	Q92JY0	Q92jy0 rhizobium m	789	6	3.4	299	16	Q98GB2	Q98gb2 rhizobium l
717	6	3.4	277	17	Q8TW77	Q8tw77 methanopyru	790	6	3.4	299	16	Q9RW02	Q9rw02 deinococcus
718	6	3.4	278	11	Q8VEC7	Q8vec7 mus musculu	791	6	3.4	300	10	Q9LW35	Q9lw35 arabidopsis
719	6	3.4	279	4	Q8WYK3	Q8wyk3 homo sapien	792	6	3.4	300	16	Q8A4B0	Q8a4b0 bacteroides
720	6	3.4	279	17	Q9YBP1	Q9ybp1 aeropyrum p	793	6	3.4	300	16	Q88IY4	Q88iy4 pseudomonas
721	6	3.4	279	17	Q8ZTB2	Q8ztb2 pyrobaculum	794	6	3.4	300	16	Q7UXX8	Q7uxx8 rhodopirell
722	6	3.4	280	2	Q9EZD1	Q9ezd1 rhizobium m	795	6	3.4	301	10	Q84ZT6	Q84zt6 asparagus o
723	6	3.4	280	4	Q8NC01	Q8nc01 homo sapien	796	6	3.4	301	16	Q89HW9	Q89hw9 bradyrhizob
724	6	3.4	280	4	Q9NZH3	Q9nzh3 homo sapien	797	6	3.4	301	16	Q887K7	Q887k7 pseudomonas
725	6	3.4	280	4	Q8IUW7	Q8iuw7 homo sapien	798	6	3.4	301	17	Q8PY26	Q8py26 methanosarc
726	6	3.4	280	11	Q9JKF8	Q9jkf8 rattus norv	799	6	3.4	302	2	Q9F8T5	Q9fst5 streptomyce
727	6	3.4	280	11	O35626	O35626 mus musculu	800	6	3.4	302	12	Q80NX1	Q80nx1 human polio
728	6	3.4	280	16	Q98I98	Q98i98 rhizobium l	801	6	3.4	302	16	Q55655	Q55655 synecocyst
729	6	3.4	280	16	Q8FPT5	Q8fpt5 corynebacte	802	6	3.4	302	16	Q884P8	Q884p8 pseudomonas
730	6	3.4	280	16	Q87PM4	Q87pm4 vibrio para	803	6	3.4	303	16	Q986A7	Q986a7 rhizobium l
731	6	3.4	281	4	Q9Y272	Q9y272 homo sapien	804	6	3.4	303	16	Q9L1P3	Q9l1p3 streptomyce
732	6	3.4	281	4	Q9NYB4	Q9nyb4 homo sapien	805	6	3.4	304	16	Q98JX7	Q98jx7 rhizobium l
733	6	3.4	281	16	Q837U8	Q837u8 enterococcu	806	6	3.4	304	16	Q92L22	Q92l22 rhizobium m
734	6	3.4	282	4	Q8N8Q3	Q8n8q3 homo sapien	807	6	3.4	305	2	Q9LAV6	Q9lav6 thermomonos
735	6	3.4	282	16	Q8PL91	Q8pl91 xanthomonas	808	6	3.4	305	10	Q8S163	Q8s163 oryza sativ
736	6	3.4	283	12	Q9DHC2	Q9dhc2 meleagrid h	809	6	3.4	305	11	Q7TQR6	Q7tqr6 mus musculu
737	6	3.4	284	4	Q86TZ2	Q86tz2 homo sapien	810	6	3.4	305	12	Q9YTK7	Q9ytk7 ateline her
738	6	3.4	284	16	Q9L4I3	Q9l4i3 salmonella	811	6	3.4	305	16	Q82G32	Q82g32 streptomyce
739	6	3.4	284	16	Q8ZBI8	Q8zbi8 yersinia pe	812	6	3.4	306	11	Q812A3	Q812a3 mus musculu
740	6	3.4	284	16	Q87Z35	Q87z35 pseudomonas	813	6	3.4	306	11	Q7TQR5	Q7tqr5 mus musculu
741	6	3.4	284	16	Q8Z9F4	Q8z9f4 salmonella	814	6	3.4	306	16	Q92RN7	Q92rn7 rhizobium m
742	6	3.4	286	11	Q8BI07	Q8bi07 mus musculu	815	6	3.4	306	16	Q8PRA1	Q8pra1 xanthomonas
743	6	3.4	286	13	Q7ZWA7	Q7zwa7 brachydanio	816	6	3.4	306	16	Q7VKQ2	Q7vkq2 haemophilus
744	6	3.4	287	5	Q86PQ0	Q86pq0 sarcocystis	817	6	3.4	306	16	Q7UP34	Q7up34 rhodopirell
745	6	3.4	287	10	Q9XFC0	Q9xfc0 mesembryant	818	6	3.4	307	5	Q94253	Q94253 caenorhabdi
746	6	3.4	288	10	O81958	O81958 hordeum vul	819	6	3.4	307	11	Q8VFW2	Q8vfw2 mus musculu



820	6	3.4	307	11	Q8VFX0	Q8vfx0 mus musculu	893	6	3.4	328	10	Q8L4E6
821	6	3.4	307	11	Q7TQR2	Q7tqr2 mus musculu	894	6	3.4	328	10	Q84UA9
822	6	3.4	308	2	Q84F71	Q84f71 photorhabdu	895	6	3.4	328	11	Q8K3K2
823	6	3.4	308	11	Q8VFX4	Q8vfx4 mus musculu	896	6	3.4	328	16	Q8P680
824	6	3.4	308	11	Q8VFW1	Q8vfw1 mus musculu	897	6	3.4	328	16	Q9RPS4
825	6	3.4	308	16	Q8XBV0	Q8xbv0 escherichia	898	6	3.4	329	10	Q9ZP15
826	6	3.4	308	16	Q97NP5	Q97np5 streptococc	899	6	3.4	329	12	O40853
827	6	3.4	308	16	Q8FDL8	Q8fdl8 escherichia	900	6	3.4	329	12	Q8JRR0
828	6	3.4	308	16	Q8DNE1	Q8dne1 streptococc	901	6	3.4	329	12	O40863
829	6	3.4	308	16	Q88M77	Q88m77 pseudomonas	902	6	3.4	329	12	Q8JRR7
830	6	3.4	308	16	Q83Q62	Q83q62 shigella fl	903	6	3.4	330	10	Q9M8K2
831	6	3.4	309	2	Q9RNN7	Q9rnn7 zymomonas m	904	6	3.4	331	5	Q9XU91
832	6	3.4	309	5	Q9XTZ2	Q9xtz2 caenorhabdi	905	6	3.4	331	15	Q98ZC4
833	6	3.4	309	11	Q8VFX3	Q8vfx3 mus musculu	906	6	3.4	331	16	Q8XWS8
834	6	3.4	309	16	Q9I217	Q9i217 pseudomonas	907	6	3.4	331	16	Q8P5W6
835	6	3.4	309	16	Q9RS04	Q9rs04 deinococcus	908	6	3.4	331	16	Q82Q67
836	6	3.4	309	16	Q8Y323	Q8y323 ralstonia s	909	6	3.4	332	2	Q939Q8
837	6	3.4	309	16	O86830	O86830 streptomyce	910	6	3.4	332	5	Q95WE9
838	6	3.4	309	16	Q8EHA3	Q8eha3 shewanella	911	6	3.4	332	16	Q8DUJ0
839	6	3.4	310	2	O30683	O30683 rhodococcus	912	6	3.4	332	16	Q8DLS9
840	6	3.4	310	5	O76606	O76606 caenorhabdi	913	6	3.4	333	6	Q8MIS4
841	6	3.4	310	16	Q8ZDK6	Q8zdk6 yersinia pe	914	6	3.4	333	16	Q9AD40
842	6	3.4	310	16	Q8EIS3	Q8eis3 shewanella	915	6	3.4	333	16	Q82EP0
843	6	3.4	310	16	Q8E3U9	Q8e3u9 streptococc	916	6	3.4	334	5	Q95WF0
844	6	3.4	310	16	Q8DY84	Q8dy84 streptococc	917	6	3.4	334	10	Q942Z2
845	6	3.4	311	6	Q95JZ0	Q95jz0 macaca fasc	918	6	3.4	334	16	Q9L0W5
846	6	3.4	311	16	Q7UYU7	Q7uyu7 rhodopirell	919	6	3.4	334	16	Q884B8
847	6	3.4	312	2	Q939Q5	Q939q5 streptomyce	920	6	3.4	334	17	Q9UY56
848	6	3.4	312	11	Q8VFW8	Q8vfw8 mus musculu	921	6	3.4	335	10	Q8RXZ9
849	6	3.4	312	11	Q8VFW7	Q8vfw7 mus musculu	922	6	3.4	335	10	Q9SAK3
850	6	3.4	312	16	Q9CNG2	Q9cng2 pasteurella	923	6	3.4	335	16	Q9RRA6
851	6	3.4	312	16	Q9S1Q6	Q9slq6 streptomyce	924	6	3.4	335	16	Q82T94
852	6	3.4	313	2	Q84HP4	Q84hp4 amycolatops	925	6	3.4	336	10	Q38758
853	6	3.4	313	11	Q8VG78	Q8vg78 mus musculu	926	6	3.4	336	10	Q9SB45
854	6	3.4	314	11	Q9Z328	Q9z328 mus musculu	927	6	3.4	336	11	Q91XK9
855	6	3.4	314	11	Q7TQR4	Q7tqr4 mus musculu	928	6	3.4	336	11	Q8BNI8
856	6	3.4	314	16	Q8UDD6	Q8udd6 agrobacteri	929	6	3.4	336	11	Q80YD8
857	6	3.4	315	16	Q93J90	Q93j90 streptomyce	930	6	3.4	336	12	Q68405
858	6	3.4	315	16	Q7UG56	Q7ug56 rhodopirell	931	6	3.4	336	16	Q9KDK4
859	6	3.4	316	5	Q8I7B0	Q8i7b0 geodia cydo	932	6	3.4	337	16	Q8PFD3
860	6	3.4	316	10	Q9FLR7	Q9flr7 arabidopsis	933	6	3.4	338	2	Q7X338
861	6	3.4	316	11	Q8VGY8	Q8vgy8 mus musculu	934	6	3.4	338	10	Q84WL6
862	6	3.4	316	16	Q87WK6	Q87wk6 pseudomonas	935	6	3.4	338	16	Q8KB62
863	6	3.4	317	5	Q7YYK7	Q7yyk7 cryptospori	936	6	3.4	339	10	Q94DM4
864	6	3.4	317	10	Q9MB47	Q9mb47 citrus unsh	937	6	3.4	339	10	Q9FI73
865	6	3.4	317	11	Q8JZW1	Q8jzw1 mus musculu	938	6	3.4	339	12	Q7TLX2
866	6	3.4	317	11	Q8VFK3	Q8vfk3 mus musculu	939	6	3.4	339	16	Q8YS39
867	6	3.4	317	11	Q8VH01	Q8vh01 mus musculu	940	6	3.4	339	16	Q89LN2
868	6	3.4	318	10	Q8W2X3	Q8w2x3 oryza sativ	941	6	3.4	339	16	Q82DB6
869	6	3.4	318	17	Q9HNH1	Q9hnh1 halobacteri	942	6	3.4	340	10	Q9SUX8
870	6	3.4	319	10	Q9SXL5	Q9exl5 chlamydomon	943	6	3.4	340	16	Q9A977
871	6	3.4	319	16	Q9RV01	Q9rv01 deinococcus	944	6	3.4	340	16	Q8ZHT9
872	6	3.4	321	2	Q9ZGC3	Q9zgc3 streptomyce	945	6	3.4	340	16	Q81NX1
873	6	3.4	321	2	Q9AM33	Q9am33 desulfovibr	946	6	3.4	341	10	Q9C8F9
874	6	3.4	321	16	Q8Y039	Q8y039 ralstonia s	947	6	3.4	341	11	Q99LQ1
875	6	3.4	321	16	Q92KA1	Q92ka1 rhizobium m	948	6	3.4	341	16	Q92TS4
876	6	3.4	322	5	Q19510	Q19510 caenorhabdi	949	6	3.4	341	16	Q81BW1
877	6	3.4	322	10	Q9SSZ9	Q9ssz9 scutellaria	950	6	3.4	342	2	Q9KW94
878	6	3.4	322	16	Q8PJM0	Q8pjm0 xanthomonas	951	6	3.4	342	10	Q84W62
879	6	3.4	322	16	Q8P871	Q8p871 xanthomonas	952	6	3.4	342	16	Q8X674
880	6	3.4	323	10	O24080	O24080 medicago sa	953	6	3.4	342	16	Q837J6
881	6	3.4	323	13	Q8UW46	Q8uw46 fugu rubrip	954	6	3.4	342	16	Q9Z922
882	6	3.4	324	3	Q9HEF2	Q9hef2 neurospora	955	6	3.4	343	4	Q96AS5
883	6	3.4	325	4	Q96H09	Q96h09 homo sapien	956	6	3.4	344	3	Q02950
884	6	3.4	325	4	Q9BTF9	Q9btf9 homo sapien	957	6	3.4	344	4	Q9NS73
885	6	3.4	325	5	Q9NKP9	Q9nkp9 leishmania	958	6	3.4	344	4	Q9BS93
886	6	3.4	325	10	Q9SSZ8	Q9ssz8 scutellaria	959	6	3.4	344	10	Q8VYN8
887	6	3.4	325	16	Q98H47	Q98h47 rhizobium l	960	6	3.4	345	16	Q82NG5
888	6	3.4	326	16	Q9K765	Q9k765 bacillus ha	961	6	3.4	345	16	Q7WIH3
889	6	3.4	326	16	Q7WGL6	Q7wgl6 bordetella	962	6	3.4	345	16	Q7W6J9
890	6	3.4	326	16	Q7W5W4	Q7w5w4 bordetella	963	6	3.4	346	10	Q9LIF7
891	6	3.4	328	2	Q84IM3	Q84im3 streptomyce	964	6	3.4	346	10	Q9LZH2
892	6	3.4	328	3	Q74776	Q74776 schizosacch	965	6	3.4	347	3	Q9P973

Q814e6 oryza sativ  
Q84ua9 artemisia a  
Q8k3k2 mus musculu  
Q8p680 xanthomonas  
Q9rps4 enterococcu  
Q9zp15 trifolium r  
O40853 influenza a  
Q8jrr0 influenza a  
O40863 influenza a  
Q8jrr7 influenza a  
Q9m8k2 arabidopsis  
Q9xu91 caenorhabdi  
Q98zc4 human immun  
Q8xws8 ralstonia s  
Q8p5w6 xanthomonas  
Q82q67 streptomyce  
Q939q8 streptomyce  
Q95we9 drosophila  
Q8duj0 streptococc  
Q8dls9 synechococc  
Q8mis4 sus scrofa  
Q9ad40 streptomyce  
Q82ep0 streptomyce  
Q95wf0 drosophila  
Q942z2 oryza sativ  
Q910w5 streptomyce  
Q884b8 pseudomonas  
Q9uy56 pyrococcus  
Q8rxz9 arabidopsis  
Q9sak3 arabidopsis  
Q9rra6 deinococcus  
Q82t94 nitrosomona  
Q38758 acer pseudo  
Q9sb45 arabidopsis  
Q91xk9 mus musculu  
Q8bni8 mus musculu  
Q80yd8 mus musculu  
Q68405 human cytom  
Q9kdk4 bacillus ha  
Q8pfd3 xanthomonas  
Q7x338 uncultured  
Q84wl6 arabidopsis  
Q8kb62 chlorobium  
Q94dm4 oryza sativ  
Q9fi73 arabidopsis  
Q7tlx2 choristoneu  
Q8ys39 anabaena sp  
Q89ln2 bradyrhizob  
Q82db6 streptomyce  
Q9sux8 arabidopsis  
Q9a977 caulobacter  
Q8zht9 yersinia pe  
Q81nx1 bacillus an  
Q9c8f9 arabidopsis  
Q99lq1 mus musculu  
Q92ts4 rhizobium m  
Q81bw1 bacillus ce  
Q9kw94 agrobacteri  
Q84w62 arabidopsis  
Q8x674 escherichia  
Q837j6 enterococcu  
Q9z922 chlamydia p  
Q96as5 homo sapien  
Q02950 saccharomyc  
Q9ns73 homo sapien  
Q9bs93 homo sapien  
Q8vyn8 arabidopsis  
Q82ng5 streptomyce  
Q7wih3 bordetella  
Q7w6j9 bordetella  
Q9lif7 arabidopsis  
Q91zh2 arabidopsis  
Q9p973 trichoderma

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:17:55 ; Search time 21 Seconds  
(without alignments)  
797.016 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 174  
Sequence: 1 RGNHWVCAHTLGHNSRGFV.....SAYAASQPQTQACPPFPSS 174

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	10	5.7	530	2 JC7983	peptidoglycan reco
2	9	5.2	161	2 T28088	hypothetical prote
3	8	4.6	119	2 C72496	hypothetical prote
4	8	4.6	173	2 S62349	L71-3 protein - fr
5	7	4.0	106	2 I52306	DNA-binding protei
6	7	4.0	123	2 AH2525	hypothetical prote
7	7	4.0	127	1 VXECSE	preprotein translo
8	7	4.0	127	2 H91241	preprotein translo
9	7	4.0	127	2 E86089	preprotein translo
10	7	4.0	127	2 AI0456	preprotein translo
11	7	4.0	127	2 AC0934	preprotein translo
12	7	4.0	127	2 T21302	hypothetical prote
13	7	4.0	144	2 T25028	hypothetical prote
14	7	4.0	215	2 T36448	probable methyltra
15	7	4.0	219	2 T18797	hypothetical prote
16	7	4.0	350	2 C91014	hypothetical prote
17	7	4.0	350	2 E85858	probable subunit o
18	7	4.0	350	2 H64988	cytochrome c-type
19	7	4.0	358	2 AB2891	hypothetical prote
20	7	4.0	358	2 F97666	molybdenum transpo
21	7	4.0	369	1 C70666	probable membrane-
22	7	4.0	376	2 B75380	probable undecapre
23	7	4.0	385	2 A91006	probable transport
24	7	4.0	385	2 B85850	probable transport
25	7	4.0	385	2 A64981	yehY protein - Esc
26	7	4.0	391	2 E83151	hypothetical prote
27	7	4.0	393	2 B83841	phosphopentomutase
28	7	4.0	394	2 B69619	phosphodeoxyribomu
29	7	4.0	394	2 T45672	hypothetical prote

30	7	4.0	401	1 A36961	pilin biogenesis p
31	7	4.0	423	2 F95208	ATPase, AAA family
32	7	4.0	560	2 AB2220	hypothetical prote
33	7	4.0	606	2 S13526	hydrogenase (EC 1.
34	7	4.0	608	2 S09790	hypothetical prote
35	7	4.0	617	2 T49444	lustrin A related
36	7	4.0	735	2 T39615	probable pre-mrna
37	7	4.0	948	2 G83264	hypothetical prote
38	7	4.0	1036	2 T30839	sarco/endoplasmic
39	7	4.0	1048	2 S27763	Ca2+-transporting
40	7	4.0	1054	2 T01556	Ca2+-transporting
41	7	4.0	1186	2 S61647	probable membrane
42	7	4.0	1217	2 C86159	hypothetical prote
43	7	4.0	1595	2 T31082	endo-1,4-beta-xyla
44	7	4.0	1687	2 T39072	DNA2-NAM7 helicase
45	7	4.0	2108	2 H70819	probable polyketid
46	7	4.0	2616	2 A57096	nudel protein prec
47	6	3.4	26	2 S30375	agglutinin - Japan
48	6	3.4	67	2 E82268	conserved hypothet
49	6	3.4	82	2 F87284	hypothetical prote
50	6	3.4	82	2 C69162	hypothetical prote
51	6	3.4	102	2 C97733	glutaredoxin 3 [im
52	6	3.4	104	2 D72542	hypothetical prote
53	6	3.4	111	2 S69593	hypothetical prote
54	6	3.4	112	2 F97161	small basic protei
55	6	3.4	117	2 B97216	probable transcrip
56	6	3.4	119	2 T36627	probable anti-sigm
57	6	3.4	119	2 T49715	hypothetical prote
58	6	3.4	123	2 B95328	hypothetical prote
59	6	3.4	124	2 T35691	probable transcrip
60	6	3.4	127	2 B97769	acetyltransferases
61	6	3.4	129	2 E70709	hypothetical prote
62	6	3.4	135	2 H82834	hypothetical prote
63	6	3.4	138	2 A87126	NrdI-family protei
64	6	3.4	140	2 AE3007	conserved hypothet
65	6	3.4	140	2 F86686	conserved hypothet
66	6	3.4	141	2 F65086	hypothetical prote
67	6	3.4	144	2 C75376	hypothetical prote
68	6	3.4	148	2 T12417	ferredoxin [2Fe-2S
69	6	3.4	151	2 S29970	desiccation-relate
70	6	3.4	153	2 T37149	probable oxidoredu
71	6	3.4	157	2 S58002	probable olfactory
72	6	3.4	157	2 G75560	conserved hypothet
73	6	3.4	157	2 S55956	probable membrane
74	6	3.4	160	2 A96277	hypothetical prote
75	6	3.4	160	2 AC2799	conserved hypothet
76	6	3.4	162	2 G82327	MSHA pilin protein
77	6	3.4	163	2 F84651	probable kinetecho
78	6	3.4	165	2 S54445	prepilin-like 17.3
79	6	3.4	165	2 T25855	hypothetical prote
80	6	3.4	165	2 T36924	probable alanine-r
81	6	3.4	169	2 G83075	type 4 fimbrial bi
82	6	3.4	172	2 D82642	conserved hypothet
83	6	3.4	173	2 T37018	probable small sec
84	6	3.4	173	2 G87383	acetyltransferase,
85	6	3.4	175	2 B83411	hypothetical prote
86	6	3.4	176	2 S39963	hypothetical prote
87	6	3.4	176	2 F83819	hypothetical prote
88	6	3.4	177	2 E72737	hypothetical prote
89	6	3.4	177	2 AC2722	conserved hypothet
90	6	3.4	177	2 T26468	hypothetical prote
91	6	3.4	178	2 C72248	ribosomal protein
92	6	3.4	178	2 T20599	hypothetical prote
93	6	3.4	182	2 D87086	conserved hypothet
94	6	3.4	183	2 E71326	hypothetical prote
95	6	3.4	185	2 S76706	hypothetical prote
96	6	3.4	185	2 B64394	hypothetical prote
97	6	3.4	186	1 F32354	hypothetical prote
98	6	3.4	187	2 G87521	transaldolase (pen
99	6	3.4	187	2 C81131	HNH endonuclease f
100	6	3.4	188	2 A75392	hypothetical prote
101	6	3.4	189	2 D70848	probable transcrip
102	6	3.4	191	1 R5BYL9	ribosomal protein



103 191 3.4 6 191 176 3.4 6 176  
104 193 3.4 6 177 177 3.4 6 177  
105 196 3.4 6 178 178 3.4 6 178  
106 196 3.4 6 179 179 3.4 6 179  
107 196 3.4 6 180 180 3.4 6 180  
108 197 3.4 6 181 181 3.4 6 181  
109 198 3.4 6 182 182 3.4 6 182  
110 199 3.4 6 183 183 3.4 6 183  
111 200 3.4 6 184 184 3.4 6 184  
112 201 3.4 6 185 185 3.4 6 185  
113 201 3.4 6 186 186 3.4 6 186  
114 201 3.4 6 187 187 3.4 6 187  
115 202 3.4 6 188 188 3.4 6 188  
116 202 3.4 6 189 189 3.4 6 189  
117 203 3.4 6 190 190 3.4 6 190  
118 204 3.4 6 191 191 3.4 6 191  
119 205 3.4 6 192 192 3.4 6 192  
120 211 3.4 6 193 193 3.4 6 193  
121 211 3.4 6 194 194 3.4 6 194  
122 212 3.4 6 195 195 3.4 6 195  
123 212 3.4 6 196 196 3.4 6 196  
124 213 3.4 6 197 197 3.4 6 197  
125 213 3.4 6 198 198 3.4 6 198  
126 215 3.4 6 199 199 3.4 6 199  
127 216 3.4 6 200 200 3.4 6 200  
128 219 3.4 6 201 201 3.4 6 201  
129 220 3.4 6 202 202 3.4 6 202  
130 222 3.4 6 203 203 3.4 6 203  
131 226 3.4 6 204 204 3.4 6 204  
132 227 3.4 6 205 205 3.4 6 205  
133 227 3.4 6 206 206 3.4 6 206  
134 228 3.4 6 207 207 3.4 6 207  
135 229 3.4 6 208 208 3.4 6 208  
136 230 3.4 6 209 209 3.4 6 209  
137 230 3.4 6 210 210 3.4 6 210  
138 233 3.4 6 211 211 3.4 6 211  
139 234 3.4 6 212 212 3.4 6 212  
140 235 3.4 6 213 213 3.4 6 213  
141 235 3.4 6 214 214 3.4 6 214  
142 235 3.4 6 215 215 3.4 6 215  
143 235 3.4 6 216 216 3.4 6 216  
144 235 3.4 6 217 217 3.4 6 217  
145 235 3.4 6 218 218 3.4 6 218  
146 237 3.4 6 219 219 3.4 6 219  
147 237 3.4 6 220 220 3.4 6 220  
148 238 3.4 6 221 221 3.4 6 221  
149 239 3.4 6 222 222 3.4 6 222  
150 245 3.4 6 223 223 3.4 6 223  
151 245 3.4 6 224 224 3.4 6 224  
152 249 3.4 6 225 225 3.4 6 225  
153 249 3.4 6 226 226 3.4 6 226  
154 249 3.4 6 227 227 3.4 6 227  
155 249 3.4 6 228 228 3.4 6 228  
156 250 3.4 6 229 229 3.4 6 229  
157 250 3.4 6 230 230 3.4 6 230  
158 250 3.4 6 231 231 3.4 6 231  
159 250 3.4 6 232 232 3.4 6 232  
160 251 3.4 6 233 233 3.4 6 233  
161 252 3.4 6 234 234 3.4 6 234  
162 252 3.4 6 235 235 3.4 6 235  
163 253 3.4 6 236 236 3.4 6 236  
164 254 3.4 6 237 237 3.4 6 237  
165 254 3.4 6 238 238 3.4 6 238  
166 255 3.4 6 239 239 3.4 6 239  
167 255 3.4 6 240 240 3.4 6 240  
168 255 3.4 6 241 241 3.4 6 241  
169 257 3.4 6 242 242 3.4 6 242  
170 257 3.4 6 243 243 3.4 6 243  
171 259 3.4 6 244 244 3.4 6 244  
172 260 3.4 6 245 245 3.4 6 245  
173 262 3.4 6 246 246 3.4 6 246  
174 262 3.4 6 247 247 3.4 6 247  
175 263 3.4 6 248 248 3.4 6 248

268 2 268 3.4 6 176  
269 1 269 3.4 6 177  
269 2 269 3.4 6 178  
269 2 269 3.4 6 179  
270 2 270 3.4 6 180  
270 2 270 3.4 6 181  
272 1 272 3.4 6 182  
272 2 272 3.4 6 183  
272 2 272 3.4 6 184  
275 2 275 3.4 6 185  
275 2 275 3.4 6 186  
275 2 275 3.4 6 187  
275 2 275 3.4 6 188  
276 2 276 3.4 6 189  
277 2 277 3.4 6 190  
277 2 277 3.4 6 191  
277 2 277 3.4 6 192  
278 2 278 3.4 6 193  
278 2 278 3.4 6 194  
279 2 279 3.4 6 195  
284 2 284 3.4 6 196  
284 2 284 3.4 6 197  
288 2 288 3.4 6 198  
288 2 288 3.4 6 199  
289 2 289 3.4 6 200  
290 2 290 3.4 6 201  
291 2 291 3.4 6 202  
291 2 291 3.4 6 203  
295 2 295 3.4 6 204  
297 2 297 3.4 6 205  
297 2 297 3.4 6 206  
297 2 297 3.4 6 207  
297 2 297 3.4 6 208  
298 2 298 3.4 6 209  
298 2 298 3.4 6 210  
298 2 298 3.4 6 211  
302 2 302 3.4 6 212  
302 2 302 3.4 6 213  
305 2 305 3.4 6 214  
307 2 307 3.4 6 215  
307 2 307 3.4 6 216  
308 2 308 3.4 6 217  
308 2 308 3.4 6 218  
308 2 308 3.4 6 219  
309 2 309 3.4 6 220  
309 2 309 3.4 6 221  
309 2 309 3.4 6 222  
309 2 309 3.4 6 223  
310 2 310 3.4 6 224  
310 2 310 3.4 6 225  
311 2 311 3.4 6 226  
312 2 312 3.4 6 227  
313 1 313 3.4 6 228  
314 2 314 3.4 6 229  
314 2 314 3.4 6 230  
314 2 314 3.4 6 231  
318 2 318 3.4 6 232  
318 2 318 3.4 6 233  
319 2 319 3.4 6 234  
323 2 323 3.4 6 235  
324 2 324 3.4 6 236  
326 2 326 3.4 6 237  
326 2 326 3.4 6 238  
327 2 327 3.4 6 239  
328 2 328 3.4 6 240  
329 2 329 3.4 6 241  
329 2 329 3.4 6 242  
331 2 331 3.4 6 243  
331 2 331 3.4 6 244  
331 2 331 3.4 6 245  
333 2 333 3.4 6 246  
334 2 334 3.4 6 247  
335 2 335 3.4 6 248

hypothetical prote  
polyhedrin - Euxoa  
hypothetical prote  
probable methyl-tr  
probable trpA prot  
hypothetical prote  
3',5'-cyclic-nucle  
conserved hypothet  
hypothetical prote  
PTS system, mannos  
PTS system, mannos  
myb-related protei  
hypothetical prote  
hypothetical prote  
conserved hypothet  
microtubule-associ  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable membrane-  
AmpE protein [limpo  
glycine hydroxymet  
endonuclease (EC 3  
conserved hypothet  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
T-cell surface gly  
chitinase (EC 3.2.  
hypothetical prote  
hypothetical prote  
ATP-dependent clp  
probable DNA-bind  
methylenetetrahydr  
hypothetical prote  
ribonucleoside-dip  
dnaJ protein - Syn  
protein twk-24 [im  
probable enzyme Z4  
conserved hypothet  
SpoIIJ family pro  
probable polyamine  
probable esterase/  
hypothetical prote  
serine proteinase  
lysR-family transc  
hypothetical prote  
hypothetical prote  
probable oxidoredu  
thymidylate syntha  
probable peroxidas  
transcription regu  
probable transcrip  
hypothetical prote  
succinate dehydrog  
hypothetical prote  
peroxidase (EC 1.1  
hypothetical prote  
cytosine-specific  
activator-regulato  
peroxisomal enoyl  
conserved ATP-GTP  
peroxidase (EC 1.1  
peroxidase (EC 1.1  
hypothetical prote  
hypothetical prote  
homeotic protein E  
probable ilvC prot  
ketol-acid reducto  
phosphoribosylform  
probable iron ABC

249	6	3.4	335	2	F96825	T8K14.13 [imported	322	6	3.4	385	2	E83506	probable MFS trans
250	6	3.4	335	2	G01448	homeobox protein H	323	6	3.4	389	2	JS0671	sarcosine oxidase
251	6	3.4	336	2	I49581	CD1.1 - mouse	324	6	3.4	389	2	T17601	hypothetical prote
252	6	3.4	336	2	T05538	hypothetical prote	325	6	3.4	389	2	C82637	porphyrin biosynth
253	6	3.4	336	2	A83801	ABC transporter (s	326	6	3.4	389	2	AB0778	probable permease
254	6	3.4	337	1	DEZYG3	glyceraldehyde-3-p	327	6	3.4	390	1	QOBET9	protein U133 - hum
255	6	3.4	339	2	AE2212	hypothetical prote	328	6	3.4	391	2	T10713	naringenin-chalcon
256	6	3.4	340	2	A10097	galactose operon r	329	6	3.4	392	2	T37173	probable cystathio
257	6	3.4	340	2	A87388	transcription regu	330	6	3.4	393	2	F83311	flavohemoprotein p
258	6	3.4	340	2	F85256	hypothetical prote	331	6	3.4	394	2	H75372	conserved hypothet
259	6	3.4	340	2	H70790	hypothetical prote	332	6	3.4	394	2	T34708	hypothetical prote
260	6	3.4	341	2	H96019	probable transcrip	333	6	3.4	395	2	G27341	hypothetical prote
261	6	3.4	341	2	F86483	unknown protein, 4	334	6	3.4	397	2	D83311	conserved hypothet
262	6	3.4	342	2	H72111	hypothetical prote	335	6	3.4	397	2	E97176	N-terminal domain
263	6	3.4	342	2	B86511	hypothetical prote	336	6	3.4	398	2	E86655	phosphoglycerate k
264	6	3.4	342	2	B90919	membrane-associate	337	6	3.4	398	2	H95057	phosphoglycerate k
265	6	3.4	342	2	G85767	hypothetical prote	338	6	3.4	398	2	A97927	phosphoglycerate k
266	6	3.4	343	2	T45415	ketol-acid reducto	339	6	3.4	398	2	F70695	hypothetical prote
267	6	3.4	344	2	S62004	hypothetical prote	340	6	3.4	398	2	D83506	conserved hypothet
268	6	3.4	346	2	T47411	hypothetical prote	341	6	3.4	399	1	A43685	polymerase-associa
269	6	3.4	347	2	A48655	transcription regu	342	6	3.4	399	1	C43685	polymerase-associa
270	6	3.4	347	2	AF3253	DNA-directed DNA p	343	6	3.4	399	2	AI0322	glycine betaine/L-
271	6	3.4	348	2	E64148	hypothetical prote	344	6	3.4	399	2	AE3148	glycosyltransferas
272	6	3.4	349	2	T42424	probable fatty-acy	345	6	3.4	399	2	F98139	sgdx protein limpo
273	6	3.4	349	2	G96584	hypothetical prote	346	6	3.4	401	2	T36882	hypothetical prote
274	6	3.4	350	2	F97254	probable translati	347	6	3.4	401	2	F83179	hypothetical prote
275	6	3.4	350	2	G84647	hypothetical prote	348	6	3.4	401	2	E84367	alcohol dehydrogen
276	6	3.4	351	2	A69808	H+/Ca2+ exchanger	349	6	3.4	403	2	AF3432	site-specific DNA-
277	6	3.4	351	2	C82808	lipoprotein precur	350	6	3.4	404	1	S62440	mevalonate kinase
278	6	3.4	353	2	G86298	protein F309.16 [i	351	6	3.4	405	2	AD2164	hypothetical prote
279	6	3.4	354	2	C81986	probable O-sialogl	352	6	3.4	407	2	D84169	molybdenum cofacto
280	6	3.4	354	2	C81040	O-sialoglycoprotei	353	6	3.4	407	2	F87494	hypothetical prote
281	6	3.4	355	2	AC2768	alcohol dehydrogen	354	6	3.4	408	1	G65132	hypothetical 44.6
282	6	3.4	355	2	T42541	hypothetical prote	355	6	3.4	408	2	T30587	hypothetical prote
283	6	3.4	355	2	E95843	hypothetical prote	356	6	3.4	408	2	T47498	hypothetical prote
284	6	3.4	357	2	S49166	cysteine proteinas	357	6	3.4	408	2	AF1660	aminopeptidases ho
285	6	3.4	358	1	JQ0805	hydrogenase (EC 1.	358	6	3.4	410	2	AG1288	aminopeptidases ho
286	6	3.4	359	2	T30382	hypothetical prote	359	6	3.4	410	2	D83744	hypothetical prote
287	6	3.4	360	2	S57777	cysteine proteinas	360	6	3.4	411	1	OXRTL	protein-lysine 6-o
288	6	3.4	360	2	T08122	cysteine endopepti	361	6	3.4	411	2	T01388	oxidoreductase hom
289	6	3.4	360	2	D90657	hypothetical prote	362	6	3.4	414	2	S60595	FixL protein - Rhi
290	6	3.4	360	2	D85508	hypothetical prote	363	6	3.4	417	2	A56171	carboxypeptidase A
291	6	3.4	362	1	S12581	cysteine proteinas	364	6	3.4	417	2	A64918	membrane-associate
292	6	3.4	362	2	S22502	cysteine proteinas	365	6	3.4	418	2	H95226	competence/damage-
293	6	3.4	363	2	B84565	probable spliceoso	366	6	3.4	418	2	D98091	competence induced
294	6	3.4	365	2	AH2606	transcription regu	367	6	3.4	419	2	D85334	hypothetical prote
295	6	3.4	365	2	G97388	transcription regu	368	6	3.4	419	2	T04530	hypothetical prote
296	6	3.4	367	2	AG0480	aspartate semialde	369	6	3.4	419	2	T04886	DAG protein homolo
297	6	3.4	367	2	G71178	hypothetical prote	370	6	3.4	420	2	S75514	sensory transducti
298	6	3.4	368	2	E97548	alcohol dehydrogen	371	6	3.4	420	2	AH3175	ABC transporter, m
299	6	3.4	369	2	S48109	neurotoxin type F	372	6	3.4	421	2	S71331	L-ascorbate peroxi
300	6	3.4	370	2	F72657	probable translati	373	6	3.4	421	2	T10190	L-ascorbate peroxi
301	6	3.4	372	2	T14193	L-ascorbate peroxi	374	6	3.4	423	2	AD3188	D-serine deaminase
302	6	3.4	372	2	AC2952	aspartate aminotra	375	6	3.4	424	2	S74788	hypothetical prote
303	6	3.4	372	2	B98331	probable transamin	376	6	3.4	425	2	B83412	hypothetical prote
304	6	3.4	373	2	T34699	hypothetical prote	377	6	3.4	426	2	C96804	hypothetical prote
305	6	3.4	373	2	T47488	hypothetical prote	378	6	3.4	430	2	T12282	L-ascorbate peroxi
306	6	3.4	374	1	S48157	alcohol dehydrogen	379	6	3.4	430	2	JQ1020	hypothetical prote
307	6	3.4	374	2	S18887	H+-exporting ATPas	380	6	3.4	432	1	S15203	glycine hydroxymet
308	6	3.4	375	1	DECHA1	alcohol dehydrogen	381	6	3.4	432	2	F96540	sterol delta7 redu
309	6	3.4	375	1	A35837	alcohol dehydrogen	382	6	3.4	434	2	S30334	glycine hydroxymet
310	6	3.4	377	2	AB2930	conserved hypothet	383	6	3.4	434	2	A43252	probable transcrip
311	6	3.4	379	2	D70846	probable metaA prot	384	6	3.4	436	2	B70510	probable acyltrans
312	6	3.4	379	2	AF2409	mannosyl transfera	385	6	3.4	437	2	AB2854	MFS permease limpo
313	6	3.4	381	2	F87553	aminotransferase,	386	6	3.4	437	2	G97630	probable sugar tra
314	6	3.4	381	2	C49757	polyglycerol phosp	387	6	3.4	438	2	A71323	hypothetical prote
315	6	3.4	381	2	F98232	probable ATP-bindi	388	6	3.4	440	2	G82426	probable phosphogl
316	6	3.4	382	2	T45301	homoserine O-acety	389	6	3.4	443	2	T48708	hypothetical prote
317	6	3.4	383	2	H84156	aminotransferase r	390	6	3.4	444	2	T47114	probable 3-carboxy
318	6	3.4	383	2	H86232	hypothetical prote	391	6	3.4	448	2	D71566	probable signal re
319	6	3.4	385	1	WZWC6E	pectate lyase (EC	392	6	3.4	450	2	E96738	hypothetical prote
320	6	3.4	385	2	I40375	N-acetylornithine	393	6	3.4	451	2	T41129	hypothetical prote
321	6	3.4	385	2	D98352	hypothetical prote	394	6	3.4	452	2	T36042	probable plasmid r



395 6 3.4 452 2 T39566 hypothetical prote  
396 6 3.4 455 2 H71534 replication initia  
397 6 3.4 455 2 G81689 chromosomal replic  
398 6 3.4 457 2 D82961 probable metallopr  
399 6 3.4 458 2 E87128 L-serine dehydrata  
400 6 3.4 461 2 E75304 hypothetical prote  
401 6 3.4 462 2 B75306 probable cycloprop  
402 6 3.4 464 1 S37616 transcription fact  
403 6 3.4 464 2 T39699 glutathione-disulf  
404 6 3.4 466 2 T41375 probable phosphogl  
405 6 3.4 467 2 C82757 glutamyl-tRNA synt  
406 6 3.4 468 2 S75389 probable phenylala  
407 6 3.4 470 2 A84772 probable glycogeni  
408 6 3.4 471 2 T43830 cytochrome-c oxida  
409 6 3.4 471 2 F82223 sodium/dicarboxyla  
410 6 3.4 472 2 JC7626 amino acid transpo  
411 6 3.4 472 2 T50637 probable histidine  
412 6 3.4 472 2 AH3353 serine-type D-Ala-  
413 6 3.4 473 2 C70925 probable amC prot  
414 6 3.4 473 2 A95997 probable amino aci  
415 6 3.4 475 1 A69149 O-antigen transpor  
416 6 3.4 476 1 A29881 ubiquinol-cytochro  
417 6 3.4 477 2 T04916 hypothetical prote  
418 6 3.4 477 2 T46917 hypothetical prote  
419 6 3.4 481 2 S76820 hypothetical prote  
420 6 3.4 483 2 S61975 glutathione-disulf  
421 6 3.4 484 2 T36427 probable rhamnose  
422 6 3.4 484 2 B98060 hypothetical prote  
423 6 3.4 486 2 B84914 hypothetical prote  
424 6 3.4 487 2 A84388 tryptophanase [imp  
425 6 3.4 488 2 T45964 LAX1 / AUX1-like p  
426 6 3.4 488 2 T32754 hypothetical prote  
427 6 3.4 489 2 F65125 probable general s  
428 6 3.4 491 2 D84241 hypothetical prote  
429 6 3.4 491 2 T36481 probable acyl-CoA  
430 6 3.4 495 2 S76889 hypothetical prote  
431 6 3.4 499 2 AH0485 probable insulinas  
432 6 3.4 502 2 B81863 exopolphosphatase  
433 6 3.4 502 2 F81080 exopolphosphatase  
434 6 3.4 503 2 AB2734 NADH dehydrogenase  
435 6 3.4 503 2 D97515 NADH dehydrogenase  
436 6 3.4 503 2 T47496 hypothetical prote  
437 6 3.4 505 2 T34924 ABC transporter AT  
438 6 3.4 505 2 A81806 probable two compo  
439 6 3.4 505 2 H81064 sensor histidine k  
440 6 3.4 507 2 T49519 hypothetical prote  
441 6 3.4 515 2 AD1048 conserved hypothet  
442 6 3.4 515 2 A70905 hypothetical prote  
443 6 3.4 516 2 T00791 purple acid phosph  
444 6 3.4 516 2 D96682 protein FlE22.18 [  
445 6 3.4 519 2 T39822 aspartate kinase -  
446 6 3.4 523 2 F75615 hypothetical prote  
447 6 3.4 527 1 C70558 probable ABC trans  
448 6 3.4 527 2 G85760 hypothetical prote  
449 6 3.4 527 2 H90858 hypothetical prote  
450 6 3.4 528 1 PAHUI alkaline phosphata  
451 6 3.4 528 2 T12309 permease 1 - commo  
452 6 3.4 528 2 AF2048 proteinase [import  
453 6 3.4 531 2 T20763 hypothetical prote  
454 6 3.4 532 2 S12076 alkaline phosphata  
455 6 3.4 532 2 D71267 probable UDP-N-ace  
456 6 3.4 533 2 E9827 protein F20B17.6 [  
457 6 3.4 534 2 A37483 F protein - Muraya  
458 6 3.4 535 1 PAHUA alkaline phosphata  
459 6 3.4 535 2 B84239 hypothetical prote  
460 6 3.4 540 2 A35986 esterase B1 - sout  
461 6 3.4 541 2 T49108 pectinesterase lik  
462 6 3.4 542 2 B81662 apolipoprotein N-a  
463 6 3.4 542 2 JQ1524 O-succinylhomoseri  
464 6 3.4 546 2 F84647 hypothetical prote  
465 6 3.4 547 2 G71307 probable oligopept  
466 6 3.4 547 2 H75632 Na(+)-linked D-ala  
467 6 3.4 549 2 D95979 probable amino aci

468 6 3.4 468 6 F69282 medium-chain acyl-  
469 6 3.4 469 6 D64412 hypothetical prote  
470 6 3.4 470 6 T44954 flagella-related p  
471 6 3.4 471 6 C71257 phenylalanine-tRNA  
472 6 3.4 472 6 A87368 urocanate hydratas  
473 6 3.4 473 6 C75596 ExoP-related prote  
474 6 3.4 474 6 E84102 flagellar hook-ass  
475 6 3.4 475 6 JW0089 organic cation tra  
476 6 3.4 476 6 T43188 heat shock protein  
477 6 3.4 477 6 T23279 hypothetical prote  
478 6 3.4 478 6 AH0488 urocanate hydratas  
479 6 3.4 479 6 T42695 hypothetical prote  
480 6 3.4 480 6 T38942 heat shock protein  
481 6 3.4 481 6 C86934 probable membrane  
482 6 3.4 482 6 G89123 protein K07C11.4 [  
483 6 3.4 483 6 H86313 protein F2H15.10 [  
484 6 3.4 484 6 A97635 adenine deaminase  
485 6 3.4 485 6 AC2858 adenine deaminase  
486 6 3.4 486 6 T18116 hypothetical prote  
487 6 3.4 487 6 A84681 nodulin-like prote  
488 6 3.4 488 6 D90058 hypothetical prote  
489 6 3.4 489 6 T04409 probable H+-export  
490 6 3.4 490 6 G83339 hypothetical prote  
491 6 3.4 491 6 H75196 hypothetical prote  
492 6 3.4 492 6 B70747 probable serine/th  
493 6 3.4 493 6 D57150 hydrogenase (EC 1.  
494 6 3.4 494 6 T14360 H+-exporting ATPas  
495 6 3.4 495 6 AC1510 internalin protein  
496 6 3.4 496 6 T35549 hypothetical prote  
497 6 3.4 497 6 RRBPBQ RNA-directed RNA p  
498 6 3.4 498 6 T12488 hypothetical prote  
499 6 3.4 499 6 TVCHRL transforming prote  
500 6 3.4 500 6 S10893 transforming prote  
501 6 3.4 501 6 T35513 probable long-chai  
502 6 3.4 502 6 B83977 hypothetical prote  
503 6 3.4 503 6 E81905 probable sulfite r  
504 6 3.4 504 6 E75119 hypothetical prote  
505 6 3.4 505 2 G70409 high affinity sulf  
506 6 3.4 605 2 T27529 hypothetical prote  
507 6 3.4 606 2 T27035 hypothetical prote  
508 6 3.4 607 1 PXNCV7 H+-exporting ATPas  
509 6 3.4 608 2 C72405 hydrogenase (EC 1.  
510 6 3.4 610 2 S37049 H+-exporting ATPas  
511 6 3.4 610 2 F83384 probable asparagin  
512 6 3.4 610 2 A83638 probable glutamine  
513 6 3.4 611 2 I50715 A2 isoform of vacu  
514 6 3.4 612 2 T40506 major facilitator  
515 6 3.4 614 2 G71306 hypothetical prote  
516 6 3.4 616 2 T00894 hypothetical prote  
517 6 3.4 617 2 I50716 A1 isoform of vacu  
518 6 3.4 617 2 B46091 H+-exporting ATPas  
519 6 3.4 617 2 T49535 probable multifunc  
520 6 3.4 618 2 S19659 H+-exporting ATPas  
521 6 3.4 618 2 A56807 H+-transporting tw  
522 6 3.4 618 2 B87566 voltage gated chlo  
523 6 3.4 618 2 S33044 hypothetical prote  
524 6 3.4 619 2 S25334 H+-exporting ATPas  
525 6 3.4 622 2 T01414 starch synthase (E  
526 6 3.4 622 2 A64494 hypothetical prote  
527 6 3.4 623 1 PXPZV9 H+-exporting ATPas  
528 6 3.4 623 2 S57790 H+-exporting ATPas  
529 6 3.4 623 2 E96818 hypothetical prote  
530 6 3.4 623 2 B83399 quinoprotein alcoh  
531 6 3.4 624 2 S74952 hypothetical prote  
532 6 3.4 631 2 G64874 probable membrane  
533 6 3.4 631 2 JC5803 ring finger protei  
534 6 3.4 632 2 S69670 hypothetical prote  
535 6 3.4 632 2 JC7155 brain finger prote  
536 6 3.4 636 2 A55428 ferroxidase precu  
537 6 3.4 638 2 T35947 probable cation-tr  
538 6 3.4 638 2 D84453 hypothetical prote  
539 6 3.4 638 2 T44763 conserved hypothet  
540 6 3.4 638 2 T02611 hypothetical prote

541	6	3.4	643	2	D86167	protein F21B7.27 [	614	6	3.4	859	2	T06429	lipoxigenase (EC 1
542	6	3.4	646	1	A55093	fatty acid transpo	615	6	3.4	866	2	G90646	hypothetical prote
543	6	3.4	648	2	PC4395	mucin 3 - human (f	616	6	3.4	866	2	B85075	probable athila tr
544	6	3.4	650	2	B87466	TonB-dependent rec	617	6	3.4	871	2	A84177	mismatch repair pr
545	6	3.4	650	2	T36419	hypothetical prote	618	6	3.4	871	2	AE3085	two component sens
546	6	3.4	653	1	UZADP5	terminal protein p	619	6	3.4	879	1	JDVLC	DNA-directed DNA p
547	6	3.4	654	2	A45357	proprotein convert	620	6	3.4	879	2	S49910	chloroplast outer
548	6	3.4	654	2	D84689	hypothetical prote	621	6	3.4	880	2	AF0397	probable acetyltra
549	6	3.4	656	2	T24344	hypothetical prote	622	6	3.4	880	2	D98201	hypothetical prote
550	6	3.4	658	2	T03416	traG protein - Agr	623	6	3.4	881	2	S25445	nitrate reductase
551	6	3.4	658	2	AB3243	conjugal transfer	624	6	3.4	884	2	S55651	DNA helicase-prima
552	6	3.4	660	2	H70798	probable cation-tr	625	6	3.4	886	2	A85905	hypothetical prote
553	6	3.4	666	1	D69103	DNA helicase (EC 3	626	6	3.4	886	2	G65036	hypothetical prote
554	6	3.4	667	2	S48285	probable glycine-t	627	6	3.4	886	2	AD0831	probable acyl-CoA
555	6	3.4	673	2	JC2458	dnak-type molecula	628	6	3.4	921	2	E86764	conserved hypothet
556	6	3.4	678	1	I77530	proprotein convert	629	6	3.4	928	2	E86546	polymorphic outer
557	6	3.4	684	2	G70744	hypothetical prote	630	6	3.4	928	2	B72077	polymorphic membra
558	6	3.4	685	2	E82297	c-di-GMP phosphodi	631	6	3.4	940	2	A38149	RNA-directed RNA p
559	6	3.4	686	2	T10548	hypothetical prote	632	6	3.4	940	2	AD1374	internalin protein
560	6	3.4	691	2	T44543	probable bacteriop	633	6	3.4	941	2	AF2415	two-component hybr
561	6	3.4	691	2	AD2766	conserved hypothet	634	6	3.4	944	2	AC2073	two-component sens
562	6	3.4	691	2	H97546	hypothetical 81.9K	635	6	3.4	951	2	D75377	probable proteinas
563	6	3.4	694	2	A96571	hypothetical prote	636	6	3.4	963	2	A55926	DNA binding protei
564	6	3.4	698	1	A47203	protein-glutamine	637	6	3.4	972	2	S77454	cation-transportin
565	6	3.4	702	2	T16832	hypothetical prote	638	6	3.4	974	2	A71466	probable zinc meta
566	6	3.4	705	2	T24343	hypothetical prote	639	6	3.4	974	2	T04910	hypothetical prote
567	6	3.4	705	2	S76729	hypothetical prote	640	6	3.4	975	2	C81728	metalloproteinase,
568	6	3.4	709	2	E64056	polypribonucleotide	641	6	3.4	982	2	T06576	probable protein k
569	6	3.4	723	2	A49613	enoyl-CoA hydratase	642	6	3.4	990	1	TQEC26	transposase - Esch
570	6	3.4	725	2	T08017	tetrafunctional pr	643	6	3.4	991	2	I49540	procollagen C-endo
571	6	3.4	726	2	E83712	ribonucleoside-dip	644	6	3.4	1006	2	T42731	atrophin-1 related
572	6	3.4	726	2	A47275	transferrin - cock	645	6	3.4	1013	2	JC2314	chitin synthase (E
573	6	3.4	727	2	D75122	hypothetical prote	646	6	3.4	1027	2	F82829	acriflavin resista
574	6	3.4	730	2	A23742	procollagen-lysine	647	6	3.4	1031	2	F83561	probable DNA polym
575	6	3.4	736	2	T06757	hypothetical prote	648	6	3.4	1034	2	A24925	beta-galactosidase
576	6	3.4	741	2	D95966	hypothetical prote	649	6	3.4	1042	2	H70203	isoleucine-tRNA li
577	6	3.4	742	2	C83061	probable aldehyde	650	6	3.4	1042	2	A57534	mucin 5AC (clone L
578	6	3.4	743	1	G02270	probable TonB-depe	651	6	3.4	1043	2	T43502	1-phosphatidylinos
579	6	3.4	746	2	AF1879	alpha-N-acetylgluc	652	6	3.4	1046	2	S67786	hypothetical prote
580	6	3.4	747	1	WZBE45	ATP-binding protei	653	6	3.4	1056	2	A53767	mucin MUC5B, trach
581	6	3.4	751	2	T46517	gene 45 protein -	654	6	3.4	1061	2	F86211	hypothetical prote
582	6	3.4	755	2	AG3069	hypothetical prote	655	6	3.4	1061	2	E86213	hypothetical prote
583	6	3.4	755	2	C98217	phosphoenolpyruvat	656	6	3.4	1062	2	T14151	Inv protein - mous
584	6	3.4	758	2	H86764	enzyme I-ntr prote	657	6	3.4	1062	2	T30255	inversin - mouse
585	6	3.4	758	2	T39210	ATP-dependent heli	658	6	3.4	1065	2	T13230	dachshund isoform
586	6	3.4	758	2	E81217	3-isopropylmalate	659	6	3.4	1068	2	E81965	probable outer mem
587	6	3.4	768	2	T45876	organic solvent to	660	6	3.4	1072	2	T13232	dachshund protein
588	6	3.4	779	2	G70926	hypothetical prote	661	6	3.4	1074	2	T13229	dachshund protein
589	6	3.4	779	2	T52332	probable fdhF prot	662	6	3.4	1081	2	T13231	dachshund protein
590	6	3.4	780	1	S39110	Ca2+-transporting	663	6	3.4	1082	2	H81020	serotype-1-specifi
591	6	3.4	781	1	TVFFDF	valosin-containing	664	6	3.4	1110	2	T19673	hypothetical prote
592	6	3.4	784	2	A86676	protein kinase Dra	665	6	3.4	1118	2	C95385	probable adenylate
593	6	3.4	790	2	T47959	carbon starvation	666	6	3.4	1125	2	S67794	probable membrane
594	6	3.4	794	2	T37989	hypothetical prote	667	6	3.4	1127	2	D70671	pyruvate carboxyla
595	6	3.4	800	2	T19627	DNA mismatch repai	668	6	3.4	1146	2	S46837	hypothetical prote
596	6	3.4	802	2	H75420	hypothetical prote	669	6	3.4	1155	2	G96539	hypothetical prote
597	6	3.4	802	2	B81794	hypothetical prote	670	6	3.4	1159	2	B95370	probable adenylate
598	6	3.4	806	2	AI2428	probable solvent t	671	6	3.4	1166	2	T29009	hypothetical prote
599	6	3.4	810	2	S75931	sucrose synthase [	672	6	3.4	1166	2	T38467	probable nuclear p
600	6	3.4	817	2	T15138	hypothetical prote	673	6	3.4	1188	2	T38467	major DNA-binding
601	6	3.4	825	2	F95963	hypothetical prote	674	6	3.4	1196	1	DNBEV1	DNA-binding protei
602	6	3.4	828	2	C88402	probable dehydroge	675	6	3.4	1196	1	DNBEHS	DNA-binding protei
603	6	3.4	831	2	G91036	protein H05C05.1 [	676	6	3.4	1196	1	DNBEHF	DNA-binding protei
604	6	3.4	831	2	A85881	probable PTS syste	677	6	3.4	1202	2	A48773	ribonuclease P (EC
605	6	3.4	831	2	D65012	hypothetical prote	678	6	3.4	1206	2	T18557	probable hydrogena
606	6	3.4	832	2	G85497	probable phosphoen	679	6	3.4	1209	2	E90994	probable regulator
607	6	3.4	835	2	T47521	probable fimbrial	680	6	3.4	1209	2	H85839	probable regulator
608	6	3.4	835	2	T15177	respiratory burst	681	6	3.4	1214	2	AG2897	conserved hypothet
609	6	3.4	840	2	D82615	hypothetical prote	682	6	3.4	1227	2	T20370	hypothetical prote
610	6	3.4	847	2	JC4836	alpha-glucuronidas	683	6	3.4	1235	2	D32433	VSG expression sit
611	6	3.4	847	2	S75975	hypothetical prote	684	6	3.4	1237	2	T45070	protein kinase hom
612	6	3.4	855	2	T07015	Cf-4A protein - to	685	6	3.4	1256	2	S60461	gene flightless-I
613	6	3.4	856	2	E75292	GGDEF family prote	686	6	3.4	1274	2	I40813	neurotoxin type F
										1321	2	JE0352	mucin MUC5B, trach



687	6	3.4	1324	2	S51622	cut3 protein - fis	760	5	2.9	34	2	I65263	homeobox protein H
688	6	3.4	1339	2	A55301	1,3-beta-D-glucan-	761	5	2.9	39	2	H82657	hypothetical prote
689	6	3.4	1345	2	T41960	major capsid prote	762	5	2.9	40	2	T05931	probable 5-methyl
690	6	3.4	1387	2	A97673	probable periplasm	763	5	2.9	40	2	S07969	T-cell receptor al
691	6	3.4	1441	2	T13889	helicase II homolo	764	5	2.9	40	2	C87340	hypothetical prote
692	6	3.4	1475	2	S42718	nuclear pore compl	765	5	2.9	40	2	B96843	protein TEL1S.1 [i
693	6	3.4	1486	2	AI0906	glutamate synthase	766	5	2.9	41	2	C71330	conserved hypothet
694	6	3.4	1517	1	F65112	glutamate synthase	767	5	2.9	42	2	I55360	ornithine-oxo-acid
695	6	3.4	1517	2	C91140	glutamate synthase	768	5	2.9	44	2	D48110	RNA recognition mo
696	6	3.4	1517	2	F85985	glutamate synthase	769	5	2.9	44	2	T29664	hypothetical prote
697	6	3.4	1538	2	AF0432	glutamate synthase	770	5	2.9	48	2	G90856	hypothetical prote
698	6	3.4	1538	2	B90924	probable ATP-depen	771	5	2.9	50	2	AD3575	hypothetical Pdlec
699	6	3.4	1538	2	F85772	ATP-dependent heli	772	5	2.9	52	4	T11818	hydroxyproline-ric
700	6	3.4	1538	2	G64922	probable ATP-depen	773	5	2.9	54	2	A38277	Ig heavy chain V r
701	6	3.4	1544	2	E59431	phosphoinositide-b	774	5	2.9	55	2	S46466	male accessory gla
702	6	3.4	1630	2	T40217	hypothetical prote	775	5	2.9	55	2	A28911	hypothetical prote
703	6	3.4	1630	2	T00390	KIAA0614 protein -	776	5	2.9	55	2	E95281	protein kinase AK1
704	6	3.4	1665	2	T29008	hypothetical prote	777	5	2.9	56	2	S66326	hypothetical prote
705	6	3.4	1689	2	JS0080	DNA-directed RNA p	778	5	2.9	56	2	A69983	hypothetical prote
706	6	3.4	1694	2	H64106	IGA-specific metal	779	5	2.9	57	2	PT0201	protein-tyrosine k
707	6	3.4	1702	2	A41859	IGA-specific metal	780	5	2.9	57	2	S66327	protein kinase AK1
708	6	3.4	1740	2	T43773	hypothetical prote	781	5	2.9	57	2	G87390	hypothetical prote
709	6	3.4	1749	2	AB1960	two-component sens	782	5	2.9	58	2	H64377	ferredoxin 2[4Fe-4
710	6	3.4	1762	2	T03222	probable polyketid	783	5	2.9	58	2	D82970	hypothetical prote
711	6	3.4	1819	2	D97132	uncharacterized ph	784	5	2.9	58	2	C82818	hypothetical prote
712	6	3.4	1820	2	A55494	latent transformin	785	5	2.9	59	2	S40961	heterogeneous ribo
713	6	3.4	1849	2	C41859	IGA-specific metal	786	5	2.9	59	2	S56139	membrane protein n
714	6	3.4	2024	2	A54103	centrosome autoant	787	5	2.9	60	2	S22148	hypothetical prote
715	6	3.4	2044	2	T13704	still life protein	788	5	2.9	60	2	A56547	sex-peptide precur
716	6	3.4	2064	2	T13707	still life protein	789	5	2.9	62	2	S01004	trypsin inhibitor
717	6	3.4	2073	2	T43311	fatty-acyl-CoA syn	790	5	2.9	62	2	AI2482	hypothetical prote
718	6	3.4	2073	2	T39207	fatty acid synthas	791	5	2.9	63	1	T1VF	proteinase inhibit
719	6	3.4	2109	2	E89066	protein H05O09.1 [	792	5	2.9	63	2	D70563	hypothetical prote
720	6	3.4	2109	2	T33247	hypothetical prote	793	5	2.9	64	2	E81389	50S ribosomal prot
721	6	3.4	2109	2	T17490	polyketide synthas	794	5	2.9	64	2	A48854	erythrocyte band 3
722	6	3.4	2121	2	T27406	hypothetical prote	795	5	2.9	65	2	A81878	hypothetical prote
723	6	3.4	2180	2	A46182	polyprotein - echo	796	5	2.9	65	2	AG2508	hypothetical prote
724	6	3.4	2254	2	T09053	low voltage-activa	797	5	2.9	66	2	PC4360	pepsinogen A - hou
725	6	3.4	2473	1	S38040	1-phosphatidylinos	798	5	2.9	66	2	S01285	hypothetical prote
726	6	3.4	2774	2	A43359	microtubule-associ	799	5	2.9	67	2	T35269	hypothetical prote
727	6	3.4	3104	2	S20473	fatty-acid synthas	800	5	2.9	68	2	JC2508	small hypothetical
728	6	3.4	3125	1	GNVSP	genome polyprotein	801	5	2.9	68	2	T44551	trypsin inhibitor
729	6	3.4	3140	1	GNVSR	genome polyprotein	802	5	2.9	68	2	S75879	hypothetical prote
730	6	3.4	3140	2	S47508	genome polyprotein	803	5	2.9	68	2	E97420	hypothetical prote
731	6	3.4	3141	1	GNVSP	genome polyprotein	804	5	2.9	69	2	B89016	protein B0213.5 [i
732	6	3.4	3172	2	S22012	erythronolide synt	805	5	2.9	69	2	T35541	hypothetical prote
733	6	3.4	3178	2	S13595	6-deoxyerythronoli	806	5	2.9	71	2	S25659	heat-stable entero
734	6	3.4	3212	2	T24692	hypothetical prote	807	5	2.9	71	2	D87036	hypothetical prote
735	6	3.4	3498	2	T22330	hypothetical prote	808	5	2.9	71	2	AB2208	hypothetical prote
736	6	3.4	3573	2	S23070	erythronolide synt	809	5	2.9	72	2	S49415	major carboxysome
737	6	3.4	3587	2	T31075	tyrocidine synthet	810	5	2.9	72	2	E89016	protein B0213.2 [i
738	6	3.4	3712	1	YGCEVC	alpha-aminoadipyl-	811	5	2.9	72	2	T25597	hypothetical prote
739	6	3.4	3828	2	T13857	trithorax protein	812	5	2.9	72	2	I49510	gene APRT protein
740	6	3.4	4574	2	G02520	plectin - human	813	5	2.9	72	2	E83733	hypothetical prote
741	6	3.4	4660	2	T42737	gp330 protein prec	814	5	2.9	72	2	E95950	hypothetical prote
742	6	3.4	4684	2	A59404	plectin [imported]	815	5	2.9	73	2	C89016	protein B0213.4 [i
743	6	3.4	4687	1	A39638	plectin - rat	816	5	2.9	73	2	JQ1921	hypothetical 8.5K
744	6	3.4	5232	2	A45086	HC-toxin synthetas	817	5	2.9	73	2	T31219	hypothetical prote
745	6	3.4	6420	2	T30283	polyketide synthas	818	5	2.9	73	2	B85629	hypothetical prote
746	6	3.4	7576	2	T17428	FK506 polyketide s	819	5	2.9	73	2	C85844	unknown protein en
747	6	3.4	9376	2	T14593	syringomycin synth	820	5	2.9	73	2	E85680	unknown protein en
748	5	2.9	15	2	B61243	dimethylaniline mo	821	5	2.9	74	2	A43921	dihydropyridine re
749	5	2.9	19	2	A61243	dimethylaniline mo	822	5	2.9	74	2	T00123	hypothetical prote
750	5	2.9	20	2	A60150	inulinase (EC 3.2.	823	5	2.9	74	2	H87490	hypothetical prote
751	5	2.9	20	2	A60293	larval serum prote	824	5	2.9	74	2	AH3104	conserved hypothet
752	5	2.9	25	2	D49253	nuclear antigen EB	825	5	2.9	75	2	D81205	hypothetical prote
753	5	2.9	25	2	PC4437	hydrogenase (EC 1.	826	5	2.9	75	2	S12191	hypothetical prote
754	5	2.9	27	2	G44636	homeotic protein H	827	5	2.9	75	2	AI2226	hypothetical prote
755	5	2.9	28	2	B56779	tetM 5'-region lea	828	5	2.9	77	2	D81212	50S ribosomal prot
756	5	2.9	31	2	E95140	hypothetical prote	829	5	2.9	77	2	H70642	probable ribosomal
757	5	2.9	32	2	E27480	cytochrome-c3 hydr	830	5	2.9	77	2	JC5052	hypothetical 8.9k
758	5	2.9	32	2	C26393	calelectrin - marb	831	5	2.9	77	2	F90702	hypothetical prote
759	5	2.9	32	2	JC5802	ovulation stimulat	832	5	2.9	77	2	A85553	hypothetical prote

833 5 2.9 77 2 AB0481 hypothetical prote 906  
834 5 2.9 77 2 G64784 ybcJ protein - Esc 907  
835 5 2.9 77 2 A53768 ubiquinol-cytochro 908  
836 5 2.9 77 2 AI3172 conjugal transfer 909  
837 5 2.9 78 1 EWSMYG cinnamycin precurs 910  
838 5 2.9 78 2 A75442 hypothetical prote 911  
839 5 2.9 78 2 E70026 hypothetical prote 912  
840 5 2.9 78 2 C71270 hypothetical prote 913  
841 5 2.9 78 2 T28368 ORF MSV207 hypothe 914  
842 5 2.9 78 2 AE0572 hypothetical secre 915  
843 5 2.9 78 2 C98182 hypothetical prote 916  
844 5 2.9 79 2 T34777 hypothetical prote 917  
845 5 2.9 79 2 G95324 probable Trad conj 918  
846 5 2.9 80 2 S04862 peroxidase (EC 1.1 919  
847 5 2.9 80 2 T45372 ribosomal protein 920  
848 5 2.9 80 2 T03197 hypothetical prote 921  
849 5 2.9 80 2 F82602 conserved hypothet 922  
850 5 2.9 80 2 C95994 hypothetical prote 923  
851 5 2.9 81 2 A87141 50S ribosomal prot 924  
852 5 2.9 81 2 T30482 zink finger protei 925  
853 5 2.9 81 2 B70727 hypothetical prote 926  
854 5 2.9 81 2 T44943 cbaC protein [impo 927  
855 5 2.9 81 2 AI3461 hypothetical prote 928  
856 5 2.9 82 2 H91246 probable DNA-bindi 929  
857 5 2.9 82 2 S15143 major coat protein 930  
858 5 2.9 82 2 C83556 coat protein B of 931  
859 5 2.9 82 2 F95849 conserved hypothet 932  
860 5 2.9 83 2 I48286 cytochrome-c oxida 933  
861 5 2.9 83 2 S13099 cytochrome-c oxida 934  
862 5 2.9 83 2 G90914 excisionase [impor 935  
863 5 2.9 83 2 T17961 hypothetical prote 936  
864 5 2.9 83 2 H81084 hypothetical prote 937  
865 5 2.9 83 2 S72951 hypothetical prote 938  
866 5 2.9 84 2 T44559 hypothetical prote 939  
867 5 2.9 84 2 D81169 hypothetical prote 940  
868 5 2.9 84 4 I59615 protein C frame sh 941  
869 5 2.9 85 1 W1BP27 gene 1.2 protein - 942  
870 5 2.9 85 2 B69770 conserved hypothet 943  
871 5 2.9 85 2 A70969 hypothetical prote 944  
872 5 2.9 86 2 D69774 hypothetical prote 945  
873 5 2.9 86 2 S00945 hypothetical prote 946  
874 5 2.9 86 2 AD1993 hypothetical prote 947  
875 5 2.9 87 1 CCYC6L cytochrome c6 - Sy 948  
876 5 2.9 87 1 W4WL51 E4 protein - human 949  
877 5 2.9 87 2 H71424 probable drought-1 950  
878 5 2.9 87 2 I69199 stem cell antigen 951  
879 5 2.9 87 2 S27040 VAR-1 protein - Pa 952  
880 5 2.9 88 2 S49416 major carboxysome 953  
881 5 2.9 88 2 D95023 hypothetical prote 954  
882 5 2.9 88 2 H86833 conserved hypothet 955  
883 5 2.9 89 2 D81959 hypothetical prote 956  
884 5 2.9 89 2 D81016 hypothetical prote 957  
885 5 2.9 89 2 T48371 hypothetical prote 958  
886 5 2.9 89 2 AG3246 conserved hypothet 959  
887 5 2.9 90 2 A49140 glutathione transf 960  
888 5 2.9 90 2 B44797 invasion plasmid a 961  
889 5 2.9 90 2 T22136 hypothetical prote 962  
890 5 2.9 91 2 T45306 hypothetical prote 963  
891 5 2.9 91 2 D70970 hypothetical prote 964  
892 5 2.9 91 2 E89881 conserved hypothet 965  
893 5 2.9 92 2 S09538 Gene 1.2 protein - 966  
894 5 2.9 92 2 AG2063 hypothetical prote 967  
895 5 2.9 92 2 E72581 hypothetical prote 968  
896 5 2.9 93 2 I84714 H2-IA-beta-2 antig 969  
897 5 2.9 93 2 D82823 hypothetical prote 970  
898 5 2.9 94 2 S49470 hypothetical prote 971  
899 5 2.9 94 2 C70564 probable PE protei 972  
900 5 2.9 94 2 H84227 hypothetical prote 973  
901 5 2.9 94 2 AB1456 hypothetical prote 974  
902 5 2.9 94 2 AC1092 hypothetical prote 975  
903 5 2.9 95 2 PQ0650 senescence-associa 976  
904 5 2.9 95 2 S72860 bacitracin resista 977  
905 5 2.9 95 2 A56842 transcription fact 978

906 5 2.9 95 2 D95201 hypothetical prote 906  
907 5 2.9 95 2 T17894 hypothetical prote 907  
908 5 2.9 95 2 S57788 hypothetical prote 908  
909 5 2.9 95 2 T08660 hypothetical prote 909  
910 5 2.9 95 2 B98068 hypothetical prote 910  
911 5 2.9 95 2 AI3378 hypothetical prote 911  
912 5 2.9 96 1 ZRBPL ren protein - phag 912  
913 5 2.9 96 1 IEECE9 hypothetical prote 913  
914 5 2.9 96 2 A85847 hypothetical prote 914  
915 5 2.9 96 2 A98002 Ren protein [impor 915  
916 5 2.9 96 2 E90873 hypothetical prote 916  
917 5 2.9 96 2 T18169 hypothetical prote 917  
918 5 2.9 96 2 T42273 hypothetical prote 918  
919 5 2.9 96 2 D81091 glutamyl-tRNA (Gln 919  
920 5 2.9 96 2 B81849 probable Glu-tRNA( 920  
921 5 2.9 97 2 T07161 late-embryogenesis 921  
922 5 2.9 97 2 T01984 late-embryogenesis 922  
923 5 2.9 97 2 E82962 hypothetical prote 923  
924 5 2.9 97 2 B96580 F15I1.5 [imported] 924  
925 5 2.9 97 2 S76575 hypothetical prote 925  
926 5 2.9 97 2 B70970 hypothetical prote 926  
927 5 2.9 97 4 S26485 cytochrome P450 21 927  
928 5 2.9 98 2 T11109 NADH2 dehydrogenas 928  
929 5 2.9 98 2 G82277 citrate (pro-3S)-l 929  
930 5 2.9 98 2 B29846 probable activator 930  
931 5 2.9 98 2 AC1295 hypothetical prote 931  
932 5 2.9 98 2 AI1666 hypothetical prote 932  
933 5 2.9 98 2 S49414 major carboxysome 933  
934 5 2.9 98 2 C88948 protein ZK1005.a [ 934  
935 5 2.9 98 2 F97294 uncharacterized pr 935  
936 5 2.9 98 2 G95849 conserved hypothet 936  
937 5 2.9 99 1 S40865 C4-dicarboxylate t 937  
938 5 2.9 99 2 D82767 integration host f 938  
939 5 2.9 99 2 B87331 ribosomal protein 939  
940 5 2.9 99 2 T34989 hypothetical prote 940  
941 5 2.9 99 2 S09522 nodC protein - Rhi 941  
942 5 2.9 99 2 T30451 probable DNA-bindi 942  
943 5 2.9 99 2 G86081 hypothetical prote 943  
944 5 2.9 99 2 G91234 hypothetical prote 944  
945 5 2.9 99 2 S60230 gibberellin-regula 945  
946 5 2.9 99 2 S60231 gibberellin-regula 946  
947 5 2.9 99 2 T10900 late-embryogenesis 947  
948 5 2.9 99 2 AE3232 conjugal transfer 948  
949 5 2.9 99 2 F91126 hypothetical prote 949  
950 5 2.9 99 2 C82022 truncated pilin NM 950  
951 5 2.9 99 2 B34955 Purkinje cell prot 951  
952 5 2.9 99 2 S21350 hypothetical prote 952  
953 5 2.9 99 2 E85971 hypothetical prote 953  
954 5 2.9 99 2 A65099 hypothetical prote 954  
955 5 2.9 99 2 AB2242 hypothetical prote 955  
956 5 2.9 100 2 PL0122 Ig heavy chain V-I 956  
957 5 2.9 100 2 C72375 hypothetical prote 957  
958 5 2.9 100 2 S20762 embryonic abundant 958  
959 5 2.9 100 2 B72664 hypothetical prote 959  
960 5 2.9 100 2 E72547 hypothetical prote 960  
961 5 2.9 101 2 S75429 hypothetical prote 961  
962 5 2.9 101 2 C72659 hypothetical prote 962  
963 5 2.9 101 2 C83073 hypothetical prote 963  
964 5 2.9 102 2 A23620 ribonuclease T1 (E 964  
965 5 2.9 102 2 G95236 PTS system, IIB co 965  
966 5 2.9 102 2 F71225 hypothetical prote 966  
967 5 2.9 102 2 C75620 hypothetical prote 967  
968 5 2.9 102 2 F91249 hypothetical prote 968  
969 5 2.9 102 2 G72700 hypothetical prote 969  
970 5 2.9 103 2 T18203 mxnA protein - Bac 970  
971 5 2.9 103 2 T49364 hypothetical prote 971  
972 5 2.9 103 2 I64116 hypothetical prote 972  
973 5 2.9 103 2 JQ2283 negatively phytoch 973  
974 5 2.9 103 2 T20347 hypothetical prote 974  
975 5 2.9 103 2 G86346 F24J8.2 protein - 975  
976 5 2.9 104 2 S56778 probable membrane 976  
977 5 2.9 104 2 S51479 drought-induced pr 977  
978 5 2.9 104 2 AD3464 hypothetical prote 978



979 5 2.9 104 2 H87548 hypothetical prote  
980 5 2.9 105 2 AG2042 thioiredoxin [impor  
981 5 2.9 105 2 JC5414 cortistatin-like p  
982 5 2.9 105 2 F71348 probable DNA-bind  
983 5 2.9 105 2 PH1526 gamma-aminobutyric  
984 5 2.9 105 2 PH1527 gamma-aminobutyric  
985 5 2.9 105 2 T09877 late embryogenesis  
986 5 2.9 105 2 T09874 conserved hypot  
987 5 2.9 105 2 E87149 NADH2 dehydrogen  
988 5 2.9 105 2 JE0384 hypothetical prote  
989 5 2.9 105 2 D82962 hypothetical prote  
990 5 2.9 105 2 T26321 urease (EC 3.5.1.5  
991 5 2.9 106 1 B36138 ferredoxin [2Fe-2S  
992 5 2.9 106 2 A22797 TRI class I histoc  
993 5 2.9 106 2 PQ0868 cell fusion protei  
994 5 2.9 106 2 PQ0871 cell fusion protei  
995 5 2.9 106 2 PQ0867 cell fusion protei  
996 5 2.9 106 2 PQ0873 cell fusion protei  
997 5 2.9 106 2 PQ0869 cell fusion protei  
998 5 2.9 106 2 PQ0866 cell fusion protei  
999 5 2.9 106 2 AG0468 frataxin-like prot  
1000 5 2.9 106 2 AG0468

ALIGNMENTS

RESULT 1  
JC7983  
peptidoglycan recognition protein, long form - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Nov-2003 #sequence\_revision 10-Nov-2003 #text\_change 24-Nov-2003  
C;Accession: JC7983  
R;Gelius, E.; Persson, C.; Karlsson, J.; Steiner, H.  
Biochem. Biophys. Res. Commun. 306, 988-994, 2003  
A;Title: A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-L-alanine a  
A;Reference number: JC7983; PMID:12821140  
A;Accession: JC7983  
A;Molecule type: mRNA  
A;Residues: 1-530 <GL>  
A;Cross-references: GB:AY282722  
C;Comment: This protein, a membrane bound protein with N-acetylmuramoyl-L-alanine amidas  
C;Genetics:  
A;Gene: pgrp-1  
C;Keywords: N-acetylmuramoyl-L-alanine amidase; peptidoglycan recognition protein

Query Match 5.7%; Score 10; DB 2; Length 530;  
Best Local Similarity 100.0%; Pred. No. 0.035;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RGWHVVGAAHT 10  
Db 441 RGWHVVGAAHT 450

RESULT 2  
T28088  
hypothetical protein ZK899.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C;Accession: T28088  
R;Kershaw, J.  
submitted to the EMBL Data Library, September 1994  
A;Reference number: Z20468  
A;Accession: T28088  
A;Molecule type: DNA  
A;Residues: 1-161 <WIL>  
A;Cross-references: EMBL:Z37140; PIDN:CAA85496.1; GSPDB:GN00028; CESP:ZK899.1  
A;Experimental source: clone ZK899  
C;Genetics:  
A;Gene: CESP:ZK899.1  
A;Map position: X

A;Introns: 46/3; 89/3; 144/2  
C;Superfamily: Caenorhabditis elegans hypothetical protein ZK899.1  
Query Match 5.2%; Score 9; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 120 TASARPPTS 128  
Db 13 TASARPPTS 21

RESULT 3  
C72496  
hypothetical protein APE2609 - Aeropyrum pernix (strain K1)  
C;Species: Aeropyrum pernix  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C;Accession: C72496  
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Take  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
DNA Res. 6, 83-101, 1999  
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aerop  
A;Reference number: A72450; MUID:99310339; PMID:10382966  
A;Accession: C72496  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-119 <KAW>  
A;Cross-references: DDBJ:AP000064; NID:G5105945; PIDN:BAA81627.1; PID:G5106316  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE2609  
C;Superfamily: Aeropyrum pernix hypothetical protein APE2609

Query Match 4.6%; Score 8; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 123 ARPPTSRR 130  
Db 48 ARPPTSRR 55

RESULT 4  
S62349  
L71-3 protein - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 19-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 24-Sep-1999  
C;Accession: S62349  
R;Wright, L.G.; Chen, T.; Thummel, C.S.; Guild, G.M.  
J. Mol. Biol. 255, 387-400, 1996  
A;Title: Molecular characterization of the 71E late puff in Drosophila melanogaster rev  
A;Reference number: S62333; MUID:96152797; PMID:8568884  
A;Accession: S62349  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-173 <WRI>  
A;Cross-references: EMBL:U24242; NID:G775230; PIDN:AAA65111.1; PID:G775231  
C;Genetics:  
A;Gene: L71-3  
A;Cross-references: FlyBase:FBgn0004590  
C;Superfamily: L71-10 protein

Query Match 4.6%; Score 8; DB 2; Length 173;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 123 ARPPTSRR 130  
Db 102 ARPPTSRR 109

RESULT 5  
I52306

DNA-binding protein - human  
C;Species: Homo sapiens (man)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jul-2000  
C;Accession: I52306  
R;Luzi, P.; Strayer, D.S.  
Biochem. Biophys. Res. Commun. 208, 153-160, 1995  
A;Title: DNA binding proteins that amplify surfactant protein B gene expression: isolation  
A;Reference number: I52306; MUID:95194400; PMID:7887923  
A;Accession: I52306  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-106 <RES>  
A;Cross-references: GB:L10403; NID:g860726; PIDN:AAA68277.1; PID:g860727

Query Match 4.0%; Score 7; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 FAGHSAG 147  
|||||  
Db 86 FAGHSAG 92

RESULT 6  
AH2525  
hypothetical protein alr7384 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120al  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AH2525  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AH2525  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-123 <KUR>  
A;Cross-references: GB:BA000020; PIDN:BAB77142.1; PID:g17134583; GSPDB:GN00180  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: alr7384  
A;Genome: plasmid

Query Match 4.0%; Score 7; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AALRTVR 40  
|||||  
Db 41 AALRTVR 47

RESULT 7  
VXECSE  
preprotein translocase secE chain - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 01-Mar-2002  
C;Accession: A35139; A32873; H65204  
R;Downing, W.L.; Sullivan, S.L.; Gottesman, M.E.; Dennis, P.P.  
J. Bacteriol. 172, 1621-1627, 1990  
A;Title: Sequence and transcriptional pattern of the essential Escherichia coli secE-nus  
A;Reference number: A35139; MUID:90170882; PMID:2137819  
A;Accession: A35139  
A;Molecule type: DNA  
A;Residues: 1-127 <DOW>  
A;Cross-references: GB:M30610; NID:g147798; PIDN:AAA24621.1; PID:g147800  
R;Schatz, P.J.; Riggs, P.D.; Jacq, A.; Fath, M.J.; Beckwith, J.  
Genes Dev. 3, 1035-1044, 1989  
A;Title: The secE gene encodes an integral membrane protein required for protein export  
A;Reference number: A32873; MUID:89378734; PMID:2673920  
A;Accession: A32873

A;Molecule type: DNA  
A;Residues: 1-127 <SCH>  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: H65204  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-127 <BLAT>  
A;Cross-references: GB:AE000472; GB:U00096; NID:g2367333; PIDN:AAC76955.1; PID:g1790413;  
A;Experimental source: strain K-12, substrain MGI655  
C;Comment: This integral inner membrane protein is an essential component of the protein  
C;Genetics:  
A;Gene: secE  
A;Map position: 90 min  
C;Superfamily: protein-export protein secE  
C;Keywords: inner membrane; protein export

Query Match 4.0%; Score 7; DB 1; Length 127;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
|||||  
Db 28 VAIVGNY 34

RESULT 8  
H91241  
preprotein translocase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C;Accession: H91241  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: H91241  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-127 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BAB38327.1; PID:g13364380; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs4904  
C;Superfamily: protein-export protein secE

Query Match 4.0%; Score 7; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
|||||  
Db 28 VAIVGNY 34

RESULT 9  
E86089  
preprotein translocase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: E86089  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: E86089  
A;Status: preliminary  
A;Molecule type: DNA



A;Residues: 1-127 <STO>  
A;Cross-references: GB:AE005174; NID:gl2518903; PIDN:AAG59177.1; GSPDB:GN00145; UWGP:Z55  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: secE  
C;Superfamily: protein-export protein secE

Query Match 4.0%; Score 7; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
| | | | |  
Db 28 VAIVGNY 34

RESULT 10  
AI0456  
preprotein translocase SecE chain [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
C;Accession: AI0456  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AI0456  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-127 <KUR>  
A;Cross-references: GB:AL590842; PIDN:CAC93221.1; PID:gl5981669; GSPDB:GN00175  
C;Genetics:  
A;Gene: secE  
C;Superfamily: protein-export protein secE

Query Match 4.0%; Score 7; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
| | | | |  
Db 28 VAIVGNY 34

RESULT 11  
AC0934  
preprotein translocase SecE chain [imported] - Salmonella enterica subsp. enterica serov  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AC0934  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AC0934  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-127 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD09493.1; PID:gl6504610; GSPDB:GN00176  
C;Genetics:  
A;Gene: STY3738  
C;Superfamily: protein-export protein secE

Query Match 4.0%; Score 7; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
| | | | |  
Db 28 VAIVGNY 34

RESULT 12  
T21302  
hypothetical protein F23B2.9 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T21302  
R;McMurray, A.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19403  
A;Accession: T21302  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-127 <WIL>  
A;Cross-references: EMBL:Z82266; PIDN:CAB05184.1; GSPDB:GN00022; CESP:F23B2.9  
A;Experimental source: clone F23B2  
C;Genetics:  
A;Gene: CESP:F23B2.9  
A;Map position: 4  
A;Introns: 19/3

Query Match 4.0%; Score 7; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 SARPPTS 128  
| | | | |  
Db 117 SARPPTS 123

RESULT 13  
T25028  
hypothetical protein T20D3.6 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T25028  
R;Lloyd, C.  
submitted to the EMBL Data Library, December 1995  
A;Reference number: Z19971  
A;Accession: T25028  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-144 <WIL>  
A;Cross-references: EMBL:Z68220; PIDN:CAA92488.1; GSPDB:GN00022; CESP:T20D3.6  
A;Experimental source: clone T20D3  
C;Genetics:  
A;Gene: CESP:T20D3.6  
A;Map position: 4  
A;Introns: 18/3; 71/1

Query Match 4.0%; Score 7; DB 2; Length 144;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 PTEAALR 37  
| | | | |  
Db 114 PTEAALR 120

RESULT 14  
T36448  
probable methyltransferase - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T36448  
R;Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, July 1999  
A;Reference number: Z21598  
A;Accession: T36448

A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-215 <SEE>  
A;Cross-references: EMBL:AL096837; PIDN:CAB48912.1; GSPDB:GN00070; SCOEDB:SCF43A.25c  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCOEDB:SCF43A.25c

Query Match 4.0%; Score 7; DB 2; Length 215;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 TARRPSV 102  
|||  
Db 186 TARRPSV 192

RESULT 15  
T18797  
hypothetical protein C01A2.6 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C;Accession: T18797  
R;Kershaw, J.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: Z19023  
A;Accession: T18797  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-219 <WIL>  
A;Cross-references: EMBL:Z81029; PIDN:CAB02701.1; GSPDB:GN00019; CESP:C01A2.6  
A;Experimental source: clone C01A2  
C;Genetics:  
A;Gene: CESP:C01A2.6  
A;Map position: 1  
A;Introns: 22/1; 160/3  
C;Superfamily: Caenorhabditis elegans hypothetical protein C01A2.6

Query Match 4.0%; Score 7; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 TSSTRPL 110  
|||  
Db 31 TSSTRPL 37

RESULT 16  
C91014  
hypothetical protein ECs3083 [imported] - Escherichia coli (strain O157:H7, substrain R1  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C;Accession: C91014  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: C91014  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-350 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BAB36506.1; PID:g13362552; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain R1MD 0509952  
C;Genetics:  
A;Gene: ECs3083

Query Match 4.0%; Score 7; DB 2; Length 350;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 RQLVRTD 69  
|||  
Db 283 RQLVRTD 289

Db 283 RQLVRTD 289

RESULT 17  
E85858  
probable subunit of heme lyase ccmH [imported] - Escherichia coli (strain O157:H7, subst  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: E85858  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhev  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: E85858  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-350 <STO>  
A;Cross-references: GB:AE005174; NID:g12516526; PIDN:AAG57329.1; GSPDB:GN00145; UWGP:Z34  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: ccmH

Query Match 4.0%; Score 7; DB 2; Length 350;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 RQLVRTD 69  
|||  
Db 283 RQLVRTD 289

RESULT 18  
H64988  
cytochrome c-type biogenesis protein CcmH precursor - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C;Accession: H64988  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cc  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: H64988  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-350 <BLAT>  
A;Cross-references: GB:AE000309; GB:U00096; NID:g1788520; PIDN:AAC75254.1; PID:g1788522;  
A;Experimental source: strain K-12, substrain MGL655  
C;Genetics:  
A;Gene: ccmH

Query Match 4.0%; Score 7; DB 2; Length 350;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 RQLVRTD 69  
|||  
Db 283 RQLVRTD 289

RESULT 19  
AB2891  
hypothetical protein modc [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C;Accession: AB2891  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.



A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AB2891  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-358 <KUR>  
A;Cross-references: GB:AE008688; PIDN:AAL43544.1; PID:g17741056; GSPDB:GN00186  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: modC  
A;Map position: circular chromosome

Query Match 4.0%; Score 7; DB 2; Length 358;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AGLLRPD 56  
|||||  
Db 45 AGLLRPD 51

RESULT 20  
F97666  
molybdenum transport protein modC (PA1861) [imported] - Agrobacterium tumefaciens (strain  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
C;Accession: F97666  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: F97666  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-358 <KUR>  
A;Cross-references: GB:AE007869; PIDN:AAK88287.1; PID:g15157755; GSPDB:GN00169  
C;Genetics:  
A;Gene: AGR\_C\_4645  
A;Map position: circular chromosome

Query Match 4.0%; Score 7; DB 2; Length 358;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AGLLRPD 56  
|||||  
Db 45 AGLLRPD 51

RESULT 21  
C70666  
probable membrane-bound ABC transporter modC - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jul-2002  
C;Accession: C70666  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: C70666  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-369 <COL>  
A;Cross-references: GB:Z83859; GB:AL123456; NID:g3261678; PIDN:CAB06128.1; PID:g1781188  
A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: modC  
C;Superfamily: molybdenum transport protein modC; ATP-binding cassette homology  
C;Keywords: ATP

F;14-215/Domain: ATP-binding cassette homology <ABC>

Query Match 4.0%; Score 7; DB 1; Length 369;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AGLLRPD 56  
|||||  
Db 45 AGLLRPD 51

RESULT 22  
B75380  
probable undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase - Deinococcus rad  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Feb-2003  
C;Accession: B75380  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; N  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A;Reference number: A75250; MUID:20036896; PMID:10567266  
A;Accession: B75380  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-376 <WHI>  
A;Cross-references: GB:AE002000; GB:AE00513; NID:g6459325; PIDN:AAF11126.1; PID:g64593  
A;Experimental source: strain R1  
C;Genetics:  
A;Gene: DR1562  
A;Map position: 1

Query Match 4.0%; Score 7; DB 2; Length 376;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 ARTASAR 124  
|||||  
Db 311 ARTASAR 317

RESULT 23  
A91006  
probable transport system permease protein ECs3017 [imported] - Escherichia coli (strai  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C;Accession: A91006  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and ger  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: A91006  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-385 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BA836440.1; PID:g13362486; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs3017

Query Match 4.0%; Score 7; DB 2; Length 385;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 DALFDLL 79  
|||||  
Db 369 DALFDLL 375

RESULT 24  
B85850

probable transport system permease protein yehY [imported] - Escherichia coli (strain O1  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: B85850  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: B85850  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-385 <STO>  
A;Cross-references: GB:AE005174; NID:g12516435; PIDN:AAG57262.1; GSPDB:GN00145; UWGP:Z33  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: yehY

Query Match 4.0%; Score 7; DB 2; Length 385;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 73 DALFDLL 79  
| | | | |  
Db 369 DALFDLL 375

RESULT 25  
A64981  
yehY protein - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C;Accession: A64981  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: A64981  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-385 <BLAT>  
A;Cross-references: GB:AE000302; GB:U00096; NID:g1788447; PIDN:AAC75191.1; PID:g1788451;  
A;Experimental source: strain K-12, substrain MG1655  
C;Genetics:  
A;Gene: yehY

Query Match 4.0%; Score 7; DB 2; Length 385;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 73 DALFDLL 79  
| | | | |  
Db 369 DALFDLL 375

RESULT 26  
B69619  
phosphodeoxyribomutase drm - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C;Accession: B69619  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: B69619  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-394 <KUN>  
A;Cross-references: GB:Z99116; GE:AL009126; NID:g2634723; PIDN:CAB14282.1; PID:g2634785  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: drm  
C;Superfamily: phosphopentomutase

Query Match 4.0%; Score 7; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAHTLGH 13  
| | | | |  
Db 32 GAHTLGH 38

RESULT 27  
B83841  
phosphopentomutase BH1530 [imported] - Bacillus halodurans (strain C-125)  
C;Species: Bacillus halodurans  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C;Accession: B83841  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: B83841  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-393 <STO>  
A;Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BAB05249.1; GSPDB:GN0(  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: BH1530  
C;Superfamily: phosphopentomutase

Query Match 4.0%; Score 7; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 48 VRAGLLR 54  
| | | | |  
Db 287 VRAGLLR 293

RESULT 28  
B69619  
phosphodeoxyribomutase drm - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C;Accession: B69619  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: B69619  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-394 <KUN>  
A;Cross-references: GB:Z99116; GE:AL009126; NID:g2634723; PIDN:CAB14282.1; PID:g2634785  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: drm  
C;Superfamily: phosphopentomutase

Query Match 4.0%; Score 7; DB 2; Length 385;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 73 DALFDLL 79  
| | | | |  
Db 369 DALFDLL 375

RESULT 26  
B83151  
hypothetical protein PA3949 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: E83151  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bu  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: E83151  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-391 <STO>  
A;Cross-references: GB:AE004813; GB:AE004091; NID:g9950134; PIDN:AAG07336.1; GSPDB:GN001



Query Match 4.0%; Score 7; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAHTLGH 13  
| | | | | | |  
Db 32 GAHTLGH 38

RESULT 29  
T45672  
hypothetical protein F14P22.110 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 28-Jul-2000  
C;Accession: T45672  
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: Z23011  
A;Accession: T45672  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-394 <DAN>  
A;Cross-references: EMBL:AL137082  
A;Experimental source: cultivar Columbia; BAC clone F14P22  
C;Genetics:  
A;Map position: 3  
A;Introns: 69/2  
A;Note: F14P22.110  
C;Superfamily: Arabidopsis thaliana hypothetical protein T20010.190

Query Match 4.0%; Score 7; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 VSLRSLH 94  
| | | | | | |  
Db 128 VSLRSLH 134

RESULT 30  
A36961  
pilin biogenesis protein pilC - Pseudomonas putida  
C;Species: Pseudomonas putida  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A36961; S35952  
R;de Groot, A.; Heijnen, I.; de Cock, H.; Filloux, A.; Tommassen, J.  
J. Bacteriol. 176, 642-650, 1994  
A;Title: Characterization of type IV pilus genes in plant growth-promoting Pseudomonas p  
A;Reference number: A36961; MUID:94131942; PMID:7905475  
A;Accession: A36961  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-401 <DEA>  
A;Cross-references: EMBL:X74276; NID:g396262; PIDN:CAAS2333.1; PID:g396264  
C;Genetics:  
A;Gene: pilC  
C;Superfamily: secretion protein xcps  
C;Keywords: transmembrane protein

Query Match 4.0%; Score 7; DB 1; Length 401;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 VRAGLLR 54  
| | | | | | |  
Db 32 VRAGLLR 38

RESULT 31  
F95208  
ATPase, AAA family [imported] - Streptococcus pneumoniae (strain TIGR4)  
C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
C;Accession: F95208  
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrisoi  
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: F95208  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-423 <KUR>  
A;Cross-references: GB:AE005672; PIDN:AAK75863.1; PID:g14973288; GSPDB:GN00164; TIGR:SI  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP1790  
C;Superfamily: Haemophilus influenzae conserved hypothetical protein HI1590

Query Match 4.0%; Score 7; DB 2; Length 423;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 SLHYTAR 98  
| | | | | | |  
Db 258 SLHYTAR 264

RESULT 32  
AB2220  
hypothetical protein all3313 [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AB2220  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ar  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: AB2220  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-560 <KUR>  
A;Cross-references: GB:BA000019; PIDN:BAB75012.1; PID:g17132408; GSPDB:GN00179  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all3313

Query Match 4.0%; Score 7; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 SSTRPLP 111  
| | | | | | |  
Db 246 SSTRPLP 252

RESULT 33  
SI3526  
hydrogenase (EC 1.18.99.1) (Fe) large chain [similarity] - Desulfovibrio vulgaris  
N;Alternate names: hydrogenase gamma  
C;Species: Desulfovibrio vulgaris  
C;Date: 23-Apr-1993 #sequence\_revision 23-Apr-1993 #text\_change 20-Apr-2001  
C;Accession: SI3526  
R;Stokkermans, J.; van Dongen, W.; Kaan, A.; van den Berg, W.; Veeger, C.  
FEMS Microbiol. Lett. 49, 217-222, 1988  
A;Title: hyd-gamma, a gene from Desulfovibrio vulgaris (Hildenborough) encodes a polype  
A;Reference number: SI3526; MUID:89306533; PMID:2663634  
A;Accession: SI3526  
A;Molecule type: DNA  
A;Residues: 1-606 <STO>  
A;Cross-references: EMBL:X57838; NID:g40827; PIDN:CAA40970.1; PID:g40828

C;Genetics:  
A;Gene: hycC  
C;Superfamily: hydrogenase (Fe) large chain; ferredoxin 2[4Fe-4S] homology  
C;Keywords: 2Fe-2S; 4Fe-4S; iron; metalloprotein; oxidoreductase  
F;33,45,48,66/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted  
F;98,102,105,111/Binding site: 4Fe-4S cluster (His, Cys, Cys) (covalent) (type N1)  
F;152,155,158,205/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted  
F;162,195,198,201/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted  
F;305,360,507,511/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted  
F;511/Binding site: diiron cofactor (Cys) #status predicted  
Query Match 4.0%; Score 7; DB 2; Length 606;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 33 EAALRTV 39  
Db 431 EAALRTV 437  
RESULT 34  
S09790  
hypothetical protein UL27 - human cytomegalovirus (strain AD169)  
C;Species: human cytomegalovirus, human herpesvirus 5  
A;Note: host Homo sapiens (man)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 08-Oct-1999  
C;Accession: S09790  
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;  
M.; Barrell, B.G.  
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
A;Reference number: S09749; MUID:90269039; PMID:2161319  
A;Accession: S09790  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-608 <CHE>  
A;Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35426.1; PID:g59632  
A;Note: this sequence was submitted to the EMBL Data Library, December 1989  
Query Match 4.0%; Score 7; DB 2; Length 608;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 47 AVRAGLL 53  
Db 69 AVRAGLL 75  
RESULT 35  
T49444  
Iustrin A related protein [imported] - Neurospora crassa  
N;Alternate names: protein B17C10.250  
C;Species: Neurospora crassa  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C;Accession: T49444  
R;Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A;Reference number: Z25022  
A;Accession: T49444  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-617 <SCH>  
A;Cross-references: EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.250  
A;Experimental source: BAC clone B17C10; strain OR74A  
C;Genetics:  
A;Gene: NCSP:B17C10.250  
A;Map position: 6  
Query Match 4.0%; Score 7; DB 2; Length 617;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 28 AALPTEA 34

Db 132 AALPTEA 138  
RESULT 36  
T39615  
probable pre-mrna splicing factor rna helicase - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Nov-2000  
C;Accession: T39615  
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.  
submitted to the EMBL Data Library, March 1998  
A;Reference number: Z21843  
A;Accession: T39615  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-735 <WOO>  
A;Cross-references: EMBL:AL022104; PIDN:CAAL7908.1; GSPDB:GN00067; SPDB:SPBC16H5.10C  
A;Experimental source: strain 972h-; cosmid c16H5  
C;Genetics:  
A;Gene: SPDB:SPBC16H5.10C  
A;Map position: 2  
Query Match 4.0%; Score 7; DB 2; Length 735;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 33 EAALRTV 39  
Db 270 EAALRTV 276  
RESULT 37  
G83264  
hypothetical protein PA3064 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: G83264  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: G83264  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-948 <STO>  
A;Cross-references: GB:AE004730; GB:AE004091; NID:g9949154; PIDN:AAG06452.1; GSPDB:GN001  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA3064  
Query Match 4.0%; Score 7; DB 2; Length 948;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 34 AALRTVR 40  
Db 820 AALRTVR 826  
RESULT 38  
T30839  
sarco/endoplasmic reticulum Ca2+-ATPase - Paramecium tetraurelia  
C;Species: Paramecium tetraurelia  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 18-Aug-2000  
C;Accession: T30839  
R;Hauser, K.; Pavlovic, N.; Kissmehl, R.; Plattner, H.  
Biochem. J. 334, 31-38, 1998  
A;Title: Molecular characterization of a sarco(endo)plasmic reticulum Ca+-ATPase gene fr  
A;Reference number: Z20897; MUID:98359725; PMID:9693098  
A;Accession: T30839



A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-1036 <HAU>  
A;Cross-references: EMBL:Y17469; PIDN:CAA76764.1  
C;Genetics:  
A;Gene: SERCA

A;Genetic code: SGC5  
A;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain homology <ATN>  
F;605-786/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 4.0%; Score 7; DB 2; Length 1036;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 LPTEAAL 36  
|||  
Db 446 LPTEAAL 452

RESULT 39  
S27763  
Ca2+-transporting ATPase (EC 3.6.3.8) LCA1 - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 19-Apr-2002  
C;Accession: A46284; S27763  
R;Wimmers, L.E.; Ewing, N.N.; Bennett, A.B.  
Proc. Natl. Acad. Sci. U.S.A. 89, 9205-9209, 1992  
A;Title: Higher plant Ca(2+)-ATPase: primary structure and regulation of mRNA abundance  
A;Reference number: A46284; MUID:93028433; PMID:1384045  
A;Accession: A46284  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 1-1048 <WI2>  
A;Cross-references: EMBL:M96324; NID:G170377; PIDN:AAA34138.1; PID:G170378  
C;Genetics:  
A;Introns: 490/1; 530/3; 623/3; 751/3; 788/2; 864/3  
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: ATP; hydrolase  
F;618-794/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 4.0%; Score 7; DB 2; Length 1048;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 LPTEAAL 36  
|||  
Db 447 LPTEAAL 453

RESULT 40  
T01556  
Ca2+-transporting ATPase (EC 3.6.3.8) ECA2 [imported] - Arabidopsis thaliana  
N;Alternate names: Ca2+-ATPase  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 19-Apr-2002  
C;Accession: T01556; T52582  
R;Dempsey, S.; Harper, M.  
submitted to the EMBL Data Library, July 1997  
A;Description: The sequence of A. thaliana TM018A10.  
A;Reference number: Z14348  
A;Accession: T01556  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1054 <DEM>  
A;Cross-references: EMBL:AF013294; NID:G2252848; PIDN:AAB62850.1; PID:G2252852; GSPDB:GN  
A;Experimental source: cultivar Columbia  
R;Pittman, J.K.; Mills, R.F.; O'Connor, C.D.; Williams, L.E.  
Gene 236, 137-147, 1999  
A;Title: Two additional type IIA Ca2+-ATPases are expressed in Arabidopsis thaliana: evi  
A;Reference number: Z26126  
A;Accession: T52582  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA

A;Residues: 1-1054 <PIT>  
A;Cross-references: EMBL:AJ132387; PIDN:CAA10659.1  
C;Genetics:  
A;Gene: ATSP:A\_TM018A10.4; ECA2  
A;Map position: 4  
A;Introns: 489/1; 529/3; 750/3; 787/2  
C;Function:

A;Description: catalyzes hydrolysis of one molecule ATP coupled to translocation of two  
es such as the sarcoplasmic or endoplasmic reticulum; Ca2+ pump  
A;Pathway: oxidative phosphorylation  
C;Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C;Keywords: ATP; calcium transport; hydrolase; ion transport; magnesium; phosphoprotein  
F;45-62/Domain: calcium binding #status predicted <CA1>  
F;115-139/Domain: calcium binding #status predicted <CA2>  
F;245-265/Domain: calcium binding #status predicted <CA3>  
F;327-346/Domain: calcium binding #status predicted <CA4>  
F;520-705/Domain: ATP binding #status predicted <ATP>  
F;617-793/Domain: ATPase nucleotide-binding domain homology <ATN>  
F;764-785/Domain: calcium binding #status predicted <CA5>  
F;368/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F;529/Binding site: ATP (Lys) #status predicted

Query Match 4.0%; Score 7; DB 2; Length 1054;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 LPTEAAL 36  
|||  
Db 449 LPTEAAL 455

Search completed: May 18, 2004, 16:22:08  
Job time : 38 secs

```

; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,865
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Villacorta, Gilberto M.
; REGISTRATION NUMBER: 34,038
; REFERENCE/DOCKET NUMBER: 4980-007-0
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-278-865-85

Query Match          3.4%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      106 STRPLP 111
Db      2 STRPLP 7

RESULT 32
US-08-278-865-88
; Sequence 88, Application US/08278865
; Patent No. 6303574
; GENERAL INFORMATION:
; APPLICANT: KAY, BRIAN K.
; APPLICANT: SPARKS, ANDREW B.
; APPLICANT: THORN, JUDITH M.
; APPLICANT: QUILLIAM, LAWRENCE A.
; APPLICANT: DER, CHANNING J.
; TITLE OF INVENTION: SRC SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,865
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Villacorta, Gilberto M.
; REGISTRATION NUMBER: 34,038
; REFERENCE/DOCKET NUMBER: 4980-007-0
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-278-865-85

Query Match          3.4%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      106 STRPLP 111
Db      2 STRPLP 7

RESULT 32
US-08-278-865-88
; Sequence 88, Application US/08278865
; Patent No. 6303574
; GENERAL INFORMATION:
; APPLICANT: KAY, BRIAN K.
; APPLICANT: SPARKS, ANDREW B.
; APPLICANT: THORN, JUDITH M.
; APPLICANT: QUILLIAM, LAWRENCE A.
; APPLICANT: DER, CHANNING J.
; TITLE OF INVENTION: SRC SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,865
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Villacorta, Gilberto M.
; REGISTRATION NUMBER: 34,038
; REFERENCE/DOCKET NUMBER: 4980-007-0
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-278-865-85

Query Match          3.4%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      106 STRPLP 111
Db      2 STRPLP 7
```

```

; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-278-865-88

Query Match          3.4%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      106 STRPLP 111
Db      2 STRPLP 7

RESULT 33
US-08-278-865-90
; Sequence 90, Application US/08278865
; Patent No. 6303574
; GENERAL INFORMATION:
; APPLICANT: KAY, BRIAN K.
; APPLICANT: SPARKS, ANDREW B.
; APPLICANT: THORN, JUDITH M.
; APPLICANT: QUILLIAM, LAWRENCE A.
; APPLICANT: DER, CHANNING J.
; TITLE OF INVENTION: SRC SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,865
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Villacorta, Gilberto M.
; REGISTRATION NUMBER: 34,038
; REFERENCE/DOCKET NUMBER: 4980-007-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-278-865-90

Query Match          3.4%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      106 STRPLP 111
Db      2 STRPLP 7
```



```
RESULT 34
US-09-500-124-82
; Sequence 82, Application US/09500124
; Patent No. 6432920
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,124
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,999
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-500-124-82

Query Match 3.4%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLP 111
Db 2 STRPLP 7

RESULT 35
US-09-500-124-85
; Sequence 85, Application US/09500124
; Patent No. 6432920
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
```

```
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,124
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,999
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-500-124-85

Query Match 3.4%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLP 111
Db 2 STRPLP 7

RESULT 36
US-09-500-124-88
; Sequence 88, Application US/09500124
; Patent No. 6432920
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/09/500,124  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/602,999  
FILING DATE: 16-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Mirock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-500-124-88

Query Match 3.4%; Score 6; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLP 111  
|||||  
Db 2 STRPLP 7

RESULT 37

US-09-205-258-365  
; Sequence 365, Application US/09205258  
; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,875  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,374  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,917  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,949  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,883  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,897  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,898  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,962  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,963  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,877  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,878  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/070,923  
; EARLIER FILING DATE: 1997-12-18  
; EARLIER APPLICATION NUMBER: 60/092,921  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1227  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 365  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (14)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-205-258-365

Query Match 3.4%; Score 6; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 VSLRSL 93  
|||||  
Db 6 VSLRSL 11

RESULT 38

US-08-602-999A-390  
; Sequence 390, Application US/08602999A  
; Patent No. 6184205



GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: KAY, Brian K.  
APPLICANT: THORN, Judith M.  
APPLICANT: QUILLIAM, Lawrence A.  
APPLICANT: DER, Channing J.  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: RIDER, James E.  
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,999A  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 390:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-602-999A-390  
Query Match 3.4%; Score 6; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 6; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
QY 108 RPLPPA 113  
Db 5 RPLPPA 10  
RESULT 39  
US-09-500-124-390  
Sequence 390, Application US/09500124  
Patent No. 6432920  
GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: KAY, Brian K.  
APPLICANT: THORN, Judith M.  
APPLICANT: QUILLIAM, Lawrence A.  
APPLICANT: DER, Channing J.  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: RIDER, James E.  
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/500,124  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: 08/602,999  
APPLICATION NUMBER: 08/602,999  
FILING DATE: 16-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 390:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-500-124-390  
Query Match 3.4%; Score 6; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 6; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
QY 108 RPLPPA 113  
Db 5 RPLPPA 10  
RESULT 40  
US-08-602-999A-411  
Sequence 411, Application US/08602999A  
Patent No. 6184205  
GENERAL INFORMATION:  
APPLICANT: SPARKS, Andrew B.  
APPLICANT: KAY, Brian K.  
APPLICANT: THORN, Judith M.  
APPLICANT: QUILLIAM, Lawrence A.  
APPLICANT: DER, Channing J.  
APPLICANT: FOWLKES, Dana M.  
APPLICANT: RIDER, James E.  
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
TITLE OF INVENTION: ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,999A  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 411:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-602-999A-411

Query Match      3.4%; Score 6; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      108 RPLPPA 113
      |||||
Db       5 RPLPPA 10
```

Search completed: May 18, 2004, 16:22:35  
Job time : 31 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:14:04 ; Search time 18 Seconds  
(without alignments)  
503.345 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 174  
Sequence: 1 RGWHVVGAAHTLGHNSRGFGV.....SAYAAQAQPTQACPPDPSS 174

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 141681 seqs, 52070155 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	50.0	576	1	PGPL_HUMAN
2	22	12.6	598	1	PGPL_PIG
3	10	5.7	530	1	PGPL_MOUSE
4	8	4.6	1658	1	ITN2_MOUSE
5	7	4.0	127	1	SECE_ECOLI
6	7	4.0	127	1	SECE_SALTY
7	7	4.0	350	1	CCMH_ECOLI
8	7	4.0	369	1	MODC_MYCTU
9	7	4.0	371	1	FLGI_RHIME
10	7	4.0	385	1	YEHY_ECOLI
11	7	4.0	393	1	DEOB_BACHD
12	7	4.0	394	1	DEOB_BACSU
13	7	4.0	401	1	PILC_PSEPU
14	7	4.0	608	1	UL27_HCMVA
15	7	4.0	735	1	DD15_SCHPO
16	7	4.0	765	1	AOC3_MOUSE
17	7	4.0	1054	1	ECA2_ARATH
18	7	4.0	1687	1	SEN1_SCHPO
19	7	4.0	2616	1	NDL_DROME
20	6	3.4	67	1	Y872_VIBCH
21	6	3.4	133	1	NXT1_DROME
22	6	3.4	138	1	NRDI_MYCLE
23	6	3.4	148	1	FER1_MESCR
24	6	3.4	150	1	RL15_THETH
25	6	3.4	151	1	DRPD_CRAPL
26	6	3.4	157	1	Y087_DEIRA
27	6	3.4	178	1	APT_STRCL
28	6	3.4	178	1	RS5_THEMA
29	6	3.4	182	1	APT_STRGB
30	6	3.4	183	1	PEGA_RHOSU
31	6	3.4	183	1	Y432_TREPA
32	6	3.4	185	1	Y754_METJA
33	6	3.4	191	1	PGHD_SHEEP
					Q96pd5 homo sapien
					Q866y3 sus scrofa
					Q8vcs0 mus musculu
					Q9z0r6 mus musculu
					P16920 escherichia
					Q919k1 salmonella
					P33925 escherichia
					P95155 mycobacteri
					Q52948 rhizobium m
					P33361 escherichia
					Q9kcn9 bacillus ha
					P46353 bacillus su
					P36641 pseudomonas
					P16763 human cytom
					O42945 schizosacch
					O70423 mus musculu
					O23087 arabidopsis
					Q92355 schizosacch
					P98159 drosophila
					Q9ktn0 vibrio chol
					Q9v3h8 drosophila
					Q9cbp9 mycobacteri
					O04683 mesembryant
					P74910 thermus the
					P22241 craterostig
					Q9ry63 deinococcus
					Q93aj8 streptomyc
					Q9xlj2 thermotoga
					Q8klq0 streptomyc
					O82868 rhodovulum
					O83447 treponema p
					Q58164 methanococc
					Q9xsm0 ovnis aries

34	6	3.4	191	1	RL9A_YEAST	P05738 saccharomyc
35	6	3.4	191	1	RL9B_YEAST	P51401 saccharomyc
36	6	3.4	196	1	AEQ1_AEQVI	P07164 aequorea vi
37	6	3.4	196	1	AEQ2_AEQVI	P02592 aequorea vi
38	6	3.4	198	1	MYTR_MITCE	P39047 mitrocoma c
39	6	3.4	205	1	HS27_HUMAN	P04792 homo sapien
40	6	3.4	212	1	TAL_BACHD	Q9k6e4 bacillus ha
41	6	3.4	212	1	TAL_BACSU	P19669 bacillus su
42	6	3.4	226	1	GSPB_AERHY	P45755 aeromonas h
43	6	3.4	227	1	Y099_SULSO	P95962 sulfolobus
44	6	3.4	228	1	UT11_ARATH	Q9m223 arabidopsis
45	6	3.4	230	1	CPSC_STRAS	Q04662 streptococc
46	6	3.4	230	1	CPSC_STRAS	P59221 streptococc
47	6	3.4	233	1	NHSR_STRAS	P52392 streptomyc
48	6	3.4	236	1	BAC3_HALVA	Q48334 haloarcula
49	6	3.4	241	1	IF3_STRAW	Q828d2 streptomyc
50	6	3.4	246	1	YL54_CORGL	P94338 corynebacte
51	6	3.4	249	1	APX1_ARATH	Q05431 arabidopsis
52	6	3.4	249	1	NPD2_PYRAE	Q8zt00 pyrobaculum
53	6	3.4	250	1	RECO_CAUCR	Q9a802 caulobacter
54	6	3.4	253	1	PYHD_CPVOP	P36701 orgyia pseu
55	6	3.4	253	1	UNG_XYLFA	Q9pa28 xylella fas
56	6	3.4	253	1	UNG_XYLFT	Q879z0 xylella fas
57	6	3.4	254	1	THIG_BIFLO	Q8g5f6 bifidobacte
58	6	3.4	254	1	YA95_MYCPN	P75597 mycoplasma
59	6	3.4	262	1	FBOL_BOMMO	P21828 bombyx mori
60	6	3.4	266	1	RHES_HUMAN	Q96d21 homo sapien
61	6	3.4	266	1	RHES_MOUSE	Q9wvd3 mus musculu
62	6	3.4	269	1	PYHD_CPVES	P36326 euxoa scand
63	6	3.4	270	1	TRPA_MYCTU	O06130 mycobacteri
64	6	3.4	278	1	PR28_MYCLE	O33075 mycobacteri
65	6	3.4	280	1	RASD_MOUSE	Q35626 mus musculu
66	6	3.4	280	1	RASD_RAT	Q9jkf8 rattus norv
67	6	3.4	281	1	APE_BRARE	O42364 brachydanio
68	6	3.4	281	1	RASD_HUMAN	Q9y272 homo sapien
69	6	3.4	281	1	TRYZ_DROER	P54630 drosophila
70	6	3.4	282	1	YUPP_MYCHO	P43051 mycoplasma
71	6	3.4	293	1	CHIE_BETVU	P36910 beta vulgar
72	6	3.4	297	1	MBL_DROME	O16011 drosophila
73	6	3.4	307	1	DNAJ_SYNY3	P73097 synechocyst
74	6	3.4	308	1	OXA2_STRPN	Q97np5 streptococc
75	6	3.4	308	1	OXA2_STRR6	Q8dnel streptococc
76	6	3.4	309	1	O5B2_HUMAN	Q96r09 homo sapien
77	6	3.4	310	1	OXA2_STRAS	Q8e3u9 streptococc
78	6	3.4	310	1	OXA2_STRAS	Q8dy84 streptococc
79	6	3.4	310	1	PR28_MYCTU	P71697 mycobacteri
80	6	3.4	312	1	TYSY_HUMAN	P04818 homo sapien
81	6	3.4	314	1	O1C1_HUMAN	Q15619 homo sapien
82	6	3.4	314	1	PE47_ARATH	Q9szb9 arabidopsis
83	6	3.4	315	1	PER4_ARATH	Q9lel5 arabidopsis
84	6	3.4	324	1	GSHB_GLOVI	Q7nf44 gloeobacter
85	6	3.4	324	1	TEST_MOUSE	Q9jhj7 mus musculu
86	6	3.4	324	1	YAGC_SCHPO	Q09875 schizosacch
87	6	3.4	326	1	BLAA_STRCI	P33651 streptomyc
88	6	3.4	327	1	ECH1_RAT	Q62651 rattus norv
89	6	3.4	327	1	OPT_CANFA	P83286 canis famil
90	6	3.4	328	1	OPT_MOUSE	Q920a0 mus musculu
91	6	3.4	329	1	PE50_ARATH	Q43731 arabidopsis
92	6	3.4	329	1	PE51_ARATH	Q9sze7 arabidopsis
93	6	3.4	331	1	HXA1_MOUSE	P09022 mus musculu
94	6	3.4	331	1	Y011_NPVOP	Q65359 orgyia pseu
95	6	3.4	333	1	HXA1_RAT	O08556 rattus norv
96	6	3.4	333	1	ILVC_MYCLE	O33114 mycobacteri
97	6	3.4	333	1	ILVC_MYCTU	O53248 mycobacteri
98	6	3.4	334	1	PUR5_PYRAB	Q9uy56 pyrococcus
99	6	3.4	335	1	HXA1_HUMAN	P49639 homo sapien
100	6	3.4	336	1	CD11_MOUSE	P11609 mus musculu
101	6	3.4	337	1	G3P_ZYMMO	P09316 zymomonas m
102	6	3.4	340	1	Y0H9_MYCTU	O69647 mycobacteri
103	6	3.4	344	1	LEU3_THEAQ	P24098 thermus aqu
104	6	3.4	347	1	ARGC_CORGL	Q59279 corynebacte
105	6	3.4	347	1	PHCA_RALSO	P52698 ralstonia s
106	6	3.4	348	1	YHHT_HAEIN	P44646 haemophilus

107	1	3.4	358	1	MBHS_AZOV1	P21950 azotobacter	180	6	3.4	467	1	SYE_XYLFA	Q9pf56 xylella fas
108	6	3.4	360	1	CYSP_HEMSP	P43156 hemerocalli	181	6	3.4	467	1	SYE_XYLFT	Q87ah7 xylella fas
109	6	3.4	362	1	AROB_HAEDU	Q7vnr5 haemophilus	182	6	3.4	468	1	SYFA_SULTO	P95961 sulfolobus
110	6	3.4	362	1	CYSP_PHAVU	P25803 phaseolus v	183	6	3.4	469	1	SYFA_SULTO	Q97ld7 sulfolobus
111	6	3.4	362	1	CYSP_VIGMU	P12412 vigna mungo	184	6	3.4	473	1	AMI3_MYCTU	Q10811 mycobacteri
112	6	3.4	366	1	DCA1_ARATH	Q96286 arabidopsis	185	6	3.4	473	1	RT4R_HUMAN	Q9bzt6 homo sapien
113	6	3.4	367	1	DCA3_BRAJU	Q9sdm8 brassica ju	186	6	3.4	475	1	GATB_THEIN	Q8rc39 thermoanaer
114	6	3.4	370	1	E2B1_AERPE	Q9ye84 aeropyrum p	187	6	3.4	475	1	SYE_BRAJA	Q89kr5 bradyrhizob
115	6	3.4	372	1	GALS_RAT	O08726 rattus norv	188	6	3.4	475	1	ZCH5_HUMAN	Q8n8u3 homo sapien
116	6	3.4	374	1	ADH1_STRCA	P80338 struthio ca	189	6	3.4	476	1	MPPB_NEUCR	P11913 neurospora
117	6	3.4	375	1	ADH1_CHICK	P23991 gallus gall	190	6	3.4	481	1	NYX_HUMAN	Q9gzus homo sapien
118	6	3.4	375	1	ADH3_COTJA	P19631 coturnix co	191	6	3.4	482	1	GATA_SYNEL	Q8dk65 synechococc
119	6	3.4	375	1	OTC_TRAHI	P78605 trametes hi	192	6	3.4	483	1	GSHR_YEAST	P41921 saccharomyc
120	6	3.4	377	1	MTB1_BRUAB	O30570 brucella ab	193	6	3.4	489	1	GSPA_ECOLI	P45756 escherichia
121	6	3.4	379	1	METX_MYCTU	O53391 mycobacteri	194	6	3.4	492	1	HSF4_MOUSE	Q9r0l1 mus musculu
122	6	3.4	381	1	DHB2_MOUSE	P51658 mus musculu	195	6	3.4	493	1	HSF4_HUMAN	Q9ulv5 homo sapien
123	6	3.4	381	1	TAGB_BACSU	P27621 bacillus su	196	6	3.4	496	1	MURE_PSEPK	Q88n81 pseudomonas
124	6	3.4	382	1	METX_MYCLE	O32874 mycobacteri	197	6	3.4	498	1	IDHP_ASPNG	P79089 aspergillus
125	6	3.4	385	1	ARGD_BACSU	P36839 bacillus su	198	6	3.4	519	1	AK_SCHPO	O60163 schizosacch
126	6	3.4	385	1	PELE_ERWCH	P04960 erwinia chr	199	6	3.4	524	1	HUNB_TRICA	Q01791 tribolium c
127	6	3.4	387	1	SSUD_RALSO	Q8xzq6 ralstonia s	200	6	3.4	528	1	PPB1_HUMAN	P09923 homo sapien
128	6	3.4	388	1	A2AB_ORYAF	O19032 orycteropus	201	6	3.4	532	1	MURD_TREPA	O83873 treponema p
129	6	3.4	388	1	MSOX_STRSQ	P40854 streptomyc	202	6	3.4	535	1	PPB1_HUMAN	P05187 homo sapien
130	6	3.4	390	1	UL33_HCMVA	P16849 human cytom	203	6	3.4	540	1	EST1_CULPI	P16854 culex pipie
131	6	3.4	391	1	CHSY_DIACA	P48389 dianthus ca	204	6	3.4	542	1	LNT_CHLMU	Q9pjk8 chlamydia m
132	6	3.4	391	1	CHSY_DIAMO	Q9lkp7 dianthus mo	205	6	3.4	542	1	MET7_NEUCR	P38675 neurospora
133	6	3.4	392	1	CYSA_STRCO	Q59829 streptomyc	206	6	3.4	548	1	Y4VI_RHISN	Q53217 rhizobium s
134	6	3.4	396	1	A2BP_MOUSE	Q9jj43 mus musculu	207	6	3.4	551	1	Y900_METJA	O58310 methanococc
135	6	3.4	396	1	EFTU_CORGL	P42439 corynebacte	208	6	3.4	553	1	SYFA_TREPA	O83938 treponema p
136	6	3.4	397	1	A2BP_HUMAN	Q9nwb1 homo sapien	209	6	3.4	554	1	HUTU_CAUCR	Q9a9m1 caulobacter
137	6	3.4	397	1	PGK_STRP3	Q8k5w7 streptococc	210	6	3.4	557	1	OCN2_HUMAN	O76082 homo sapien
138	6	3.4	397	1	PGK_STRP8	Q8nzg3 streptococc	211	6	3.4	561	1	VATA_MAIZE	P49087 zea mays (m
139	6	3.4	397	1	PGK_STRPY	P82487 streptococc	212	6	3.4	563	1	HUTU_YERPE	Q8za09 yersinia pe
140	6	3.4	398	1	PGK_LACLA	Q9ciw1 lactococcus	213	6	3.4	565	1	UB21_HUMAN	Q9uk80 homo sapien
141	6	3.4	398	1	PGK_STRA3	Q8e3f0 streptococc	214	6	3.4	566	1	UB21_MOUSE	Q9qzl6 mus musculu
142	6	3.4	398	1	PGK_STRA5	Q8dxt0 streptococc	215	6	3.4	575	1	UL87_EBV	P25215 epstein-bar
143	6	3.4	398	1	PGK_STRMU	Q8dvv2 streptococc	216	6	3.4	580	1	VATA_HORVU	Q40002 hordeum vul
144	6	3.4	398	1	PGK_STRPN	Q97s89 streptococc	217	6	3.4	584	1	PHAC_BORPE	P35077 bordetella
145	6	3.4	398	1	PGK_STRR6	Q8dqx8 streptococc	218	6	3.4	585	1	PKNI_MYCTU	Q10964 mycobacteri
146	6	3.4	399	1	RRPP_PL4HA	P22044 human parai	219	6	3.4	587	1	VATA_CYACA	P48414 cyanidium c
147	6	3.4	399	1	RRPP_PL4HB	P21738 human parai	220	6	3.4	589	1	RRPO_BPQBE	P14647 bacterioph
148	6	3.4	400	1	TRUD_METKA	Q8txj7 methanopyru	221	6	3.4	593	1	REL_CHICK	O95198 homo sapien
149	6	3.4	404	1	KIME_SCHPO	Q09780 schizosacch	222	6	3.4	598	1	VATA_NEUCR	P16236 gallus gall
150	6	3.4	405	1	Y574_HUMAN	O60320 homo sapien	223	6	3.4	607	1	Y4PA_RHISN	P11592 neurospora
151	6	3.4	408	1	YHFW_ECOLI	P45549 escherichia	224	6	3.4	609	1	VATA_TRYCO	P55610 rhizobium s
152	6	3.4	411	1	LYOX_RAT	P16636 rattus norv	225	6	3.4	610	1	VAA1_ACEAT	Q26975 trypanosoma
153	6	3.4	417	1	CBP2_HUMAN	P48052 homo sapien	226	6	3.4	613	1	VAA2_ACEAT	Q38676 acetabulari
154	6	3.4	417	1	NTR2_MOUSE	P70310 mus musculu	227	6	3.4	613	1	VAA1_DROME	Q38677 acetabulari
155	6	3.4	417	1	UIDC_ECOLI	Q47706 escherichia	228	6	3.4	614	1	VAA1_DROME	P48602 drosophila
156	6	3.4	418	1	CINA_STRPN	P54184 streptococc	229	6	3.4	614	1	VAA2_DROME	Q27331 drosophila
157	6	3.4	424	1	GLA2_RALSO	Q8xtqi ralstonia s	230	6	3.4	614	1	Y577_TREPA	O83586 treponema p
158	6	3.4	425	1	SYH_PHOLL	Q7n705 photorhabdu	231	6	3.4	615	1	VATA_AEDAE	O16109 aedes aegyp
159	6	3.4	426	1	THCD_RHOER	P43494 rhodococcus	232	6	3.4	617	1	VAA1_BOVIN	P31404 bos taurus
160	6	3.4	428	1	SP7_MOUSE	Q8vi67 mus musculu	233	6	3.4	617	1	VAA1_HUMAN	P38606 homo sapien
161	6	3.4	431	1	SP7_HUMAN	Q8tdd2 homo sapien	234	6	3.4	617	1	VAA1_MOUSE	P50516 mus musculu
162	6	3.4	432	1	GLYA_BRAJA	P24060 bradyrhizob	235	6	3.4	617	1	VAA1_PIG	Q29048 sus scrofa
163	6	3.4	432	1	ST7R_ARATH	Q9ldu6 arabidopsis	236	6	3.4	619	1	VATA_SCHPO	P31406 schizosacch
164	6	3.4	434	1	ADA2_YEAST	Q02336 saccharomyc	237	6	3.4	622	1	SYMM_NEUCR	Q9c2h9 neurospora
165	6	3.4	434	1	GLYA_HYPME	P34895 hyphomicrob	238	6	3.4	622	1	VATA_PHAAU	P13548 phaseolus a
166	6	3.4	434	1	GLYA_METEX	P50435 methyllobact	239	6	3.4	622	1	YF54_METJA	Q58949 methanococc
167	6	3.4	437	1	SPA4_HUMAN	Q9npe6 homo sapien	240	6	3.4	623	1	VATA_ARATH	Q9z4j7 pseudomonas
168	6	3.4	438	1	GLYA_BRUME	Q8ygg7 brucella me	241	6	3.4	623	1	VATA_ARATH	O23654 arabidopsis
169	6	3.4	438	1	GLYA_BRUSU	Q8glf1 brucella su	242	6	3.4	623	1	VATA_BETVU	Q39442 beta vulgar
170	6	3.4	438	1	Y456_TREPA	O83469 treponema p	243	6	3.4	623	1	VATA_BRANA	Q39291 brassica na
171	6	3.4	443	1	SPA4_MOUSE	Q9jjf2 mus musculu	244	6	3.4	623	1	VATA_CITUN	Q9sm09 citrus unsh
172	6	3.4	444	1	SPA4_RAT	O55034 rattus norv	245	6	3.4	623	1	VATA_DAUCA	P09469 daucus caro
173	6	3.4	455	1	DNA2_CHLMU	Q9pkb9 chlamydia m	246	6	3.4	623	1	VATA_GOSHI	P31405 gossypium h
174	6	3.4	455	1	DNA2_CHLTR	O84277 chlamydia t	247	6	3.4	628	1	YS9A_CAEEL	Q960a0 caenorhabdi
175	6	3.4	464	1	ELG_DROME	Q04688 drosophila	248	6	3.4	629	1	Y492_MYCTU	Q11157 mycobacteri
176	6	3.4	464	1	GSHR_SCHPO	P78965 schizosacch	249	6	3.4	631	1	YCIQ_ECOLI	P45848 escherichia
177	6	3.4	465	1	FXD3_MOUSE	Q61060 mus musculu	250	6	3.4	631	1	Z179_RAT	O70418 rattus norv
178	6	3.4	466	1	SERA_SCHPO	P87228 schizosacch	251	6	3.4	632	1	Z179_HUMAN	Q9ulx5 homo sapien
179	6	3.4	466	1	ZIC3_MOUSE	Q62521 mus musculu	252	6	3.4	636	1	FET3_YEAST	P38993 saccharomyc



253	6	3.4	640	1	TRAG_AGRU	Q44360 agrobacteri	326	6	3.4	1202	1	RPM2_YEAST	Q02773 saccharomyc
254	6	3.4	643	1	ZN23_HUMAN	P17027 homo sapien	327	6	3.4	1231	1	YKT3_CABEL	P34314 caenorhabdi
255	6	3.4	646	1	FATP_MOUSE	Q60714 mus musculu	328	6	3.4	1233	1	MUSA_HUMAN	P98088 homo sapien
256	6	3.4	646	1	FATP_RAT	P97849 rattus norv	329	6	3.4	1235	1	CYAA_TRYBB	Q26721 trypanosoma
257	6	3.4	653	1	TERM_ADE05	P04499 human aden	330	6	3.4	1256	1	FLII_DROME	Q24020 drosophila
258	6	3.4	658	1	TRAG_AGR5	Q44346 agrobacteri	331	6	3.4	1274	1	BXF_CLOBO	P30996 clostridium
259	6	3.4	660	1	YG39_PSESM	Q886e0 pseudomonas	332	6	3.4	1282	1	DOME_DROME	Q9vwe0 drosophila
260	6	3.4	667	1	SYG_YEAST	P98088 saccharomyc	333	6	3.4	1324	1	SMC4_SCHPO	P41004 schizosacch
261	6	3.4	671	1	PSAA_AMPOP	Q9mtq4 amphidinium	334	6	3.4	1345	1	VCAP_HSV7J	P52347 human herpe
262	6	3.4	678	1	DAPT_MOUSE	P98192 mus musculu	335	6	3.4	1475	1	N153_HUMAN	P49790 homo sapien
263	6	3.4	678	1	DAPT_RAT	Q9es71 rattus norv	336	6	3.4	1503	1	TRL2_HUMAN	O94759 homo sapien
264	6	3.4	678	1	PSAA_AMPCA	P58309 amphidinium	337	6	3.4	1517	1	GLTB_ECOLI	P09831 escherichia
265	6	3.4	680	1	DAPT_HUMAN	O15228 homo sapien	338	6	3.4	1538	1	LHR_ECOLI	P30015 escherichia
266	6	3.4	697	1	TGM2_CHICK	Q01841 gallus gall	339	6	3.4	1689	1	RPAL_SCHPO	P15398 schizosacch
267	6	3.4	709	1	PNP_HAEIN	P44584 haemophilus	340	6	3.4	1694	1	IGA0_HAEIN	P44969 haemophilus
268	6	3.4	713	1	GAMP_HUMAN	O75325 homo sapien	341	6	3.4	1702	1	IGA2_HAEIN	P45384 haemophilus
269	6	3.4	722	1	ECHP_HUMAN	Q08426 homo sapien	342	6	3.4	1849	1	IGA4_HAEIN	P45386 haemophilus
270	6	3.4	725	1	MEPA_BRANA	O49809 b glyoxysom	343	6	3.4	2044	1	SIF2_DROME	P91620 drosophila
271	6	3.4	726	1	TRF_BLADI	Q02942 blaberus di	344	6	3.4	2064	1	SIFI_DROME	P91621 drosophila
272	6	3.4	730	1	PLOI_CHICK	P24802 gallus gall	345	6	3.4	2073	1	FAS1_SCHPO	Q9uug0 s fatty aci
273	6	3.4	737	1	DPP3_HUMAN	Q9ny33 homo sapien	346	6	3.4	2080	1	TOXC_COCCA	Q92215 c putative
274	6	3.4	738	1	DPP3_MOUSE	Q99kk7 mus musculu	347	6	3.4	2109	1	PKS1_ASPPA	Q12053 aspergillus
275	6	3.4	738	1	DPP3_RAT	O55096 rattus norv	348	6	3.4	2180	1	POLG_EC22H	Q66578 e genome po
276	6	3.4	743	1	ANAG_HUMAN	P54802 homo sapien	349	6	3.4	2193	1	POLG_CX16G	Q65900 c genome po
277	6	3.4	745	1	RED2_MOUSE	Q9ji20 mus musculu	350	6	3.4	2193	1	POLG_CX16T	Q9qf31 c genome po
278	6	3.4	746	1	RED2_RAT	P97616 rattus norv	351	6	3.4	2254	1	CCAG_RAT	O54898 rattus norv
279	6	3.4	747	1	VTER_VZVD	P09294 varicella-z	352	6	3.4	2282	1	ZAN_RABIT	P57999 oryctolagus
280	6	3.4	749	1	DNK3_SYNP7	P50022 synechococc	353	6	3.4	2473	1	TOR2_YEAST	P32600 saccharomyc
281	6	3.4	758	1	LEU2_SCHPO	O14289 schizosacch	354	6	3.4	2774	1	MAPA_RAT	P34926 rattus norv
282	6	3.4	762	1	GLGB_NEIDE	Q9rq15 gleisseria d	355	6	3.4	2805	1	MAPA_HUMAN	P78559 homo sapien
283	6	3.4	779	1	LON_MYCSM	O31147 mycobacteri	356	6	3.4	3125	1	POLG_PPVNA	P17766 p genome po
284	6	3.4	779	1	YT00_MYCTU	Q10821 mycobacteri	357	6	3.4	3140	1	POLG_PPVRA	P17767 p genome po
285	6	3.4	780	1	AFG2_YEAST	P32794 saccharomyc	358	6	3.4	3141	1	POLG_PPVD	P13529 p genome po
286	6	3.4	781	1	KRAF_DROME	P11346 drosophila	359	6	3.4	3172	1	ERY3_SACER	Q03133 saccharopol
287	6	3.4	794	1	PMS1_SCHPO	P54280 schizosacch	360	6	3.4	3321	1	PCN2_HUMAN	Q95613 homo sapien
288	6	3.4	800	1	ARNT_RAT	P41739 rattus norv	361	6	3.4	3567	1	ERY2_SACER	Q03132 saccharopol
289	6	3.4	810	1	SYFB_SYNY3	P74296 synechocyst	362	6	3.4	3587	1	TYCB_BREPA	O30408 b tyrocidin
290	6	3.4	831	1	FRYA_ECOS7	Q8xbq8 escherichia	363	6	3.4	3712	1	ACVS_CEPAC	P25464 cephalospor
291	6	3.4	831	1	FRYA_ECOL6	Q8ffd8 escherichia	364	6	3.4	3828	1	TRX_DROVI	Q24742 drosophila
292	6	3.4	831	1	FRYA_ECOLI	P77439 escherichia	365	6	3.4	4473	1	PLEI_CRIGR	Q9ji55 cricetulus
293	6	3.4	831	1	FRYA_SHIFL	Q83qp3 shigella fl	366	6	3.4	4655	1	LRP2_HUMAN	P98164 homo sapien
294	6	3.4	847	1	AGUA_TRIRE	Q99024 trichoderma	367	6	3.4	4660	1	LRP2_RAT	P98158 rattus norv
295	6	3.4	863	1	K6A5_MOUSE	Q8c050 mus musculu	368	6	3.4	4684	1	PLE1_HUMAN	Q15149 homo sapien
296	6	3.4	879	1	DPOL_WHVI	P03160 woodchuck h	369	6	3.4	4687	1	PLE1_RAT	P30427 rattus norv
297	6	3.4	881	1	NIA1_PHAVU	P39865 phaseolus v	370	6	3.4	4835	1	MDN1_GIALA	Q8t5t1 giardia lam
298	6	3.4	885	1	PLSB_XANAC	Q8pes0 xanthomonas	371	6	3.4	5065	1	EPPL_HUMAN	P58107 homo sapien
299	6	3.4	886	1	PLSB_XANCP	Q8p3e3 xanthomonas	372	6	3.4	5217	1	HTS1_COCCA	Q01886 cochliobolu
300	6	3.4	886	1	YFIQ_ECOLI	P76594 escherichia	373	6	3.4	5703	1	MUSB_HUMAN	Q9hc84 homo sapien
301	6	3.4	903	1	RIN2_MOUSE	Q9d684 mus musculu	374	5	2.9	15	1	ALB2_TRASC	P81189 trachemys s
302	6	3.4	909	1	CSKP_RAT	Q52915 rattus norv	375	5	2.9	15	1	UC23_MAIZE	P80629 zea mays (m
303	6	3.4	924	1	CSKP_MOUSE	O70589 mus musculu	376	5	2.9	22	1	PSBQ_ORYSA	P83646 oryza sativ
304	6	3.4	926	1	CSKP_HUMAN	O14936 homo sapien	377	5	2.9	32	1	PHNS_DESMU	P13062 desulfovibr
305	6	3.4	928	1	PMP9_CHLPN	Q92398 chlamydia p	378	5	2.9	38	1	PSAI_SYNP2	Q54752 synechococc
306	6	3.4	963	1	RFX1_MOUSE	P48377 mus musculu	379	5	2.9	40	1	RL36_CORGL	Q8nmn8 corynebacte
307	6	3.4	972	1	CTAI_BACCI	P94286 bacillus ci	380	5	2.9	49	1	SSPO_BACAA	Q8ly79 bacillus an
308	6	3.4	985	1	AGLU_ASPNG	P56526 aspergillus	381	5	2.9	49	1	SSPO_BACCR	Q81af5 bacillus ce
309	6	3.4	990	1	TNP7_ECOLI	P13694 escherichia	382	5	2.9	53	1	YORU_TTV1	P19305 thermoprote
310	6	3.4	991	1	BMP1_MOUSE	P98063 mus musculu	383	5	2.9	55	1	A70A_DROMA	O18666 drosophila
311	6	3.4	996	1	VGNM_RCMV	P13561 red clover	384	5	2.9	55	1	A70A_DROME	P05623 drosophila
312	6	3.4	1013	1	CHSA_EMENI	P30584 emericeella	385	5	2.9	55	1	A70A_DROSE	O18417 drosophila
313	6	3.4	1034	1	BGAL_KLEPN	P06219 klebsiella	386	5	2.9	55	1	MERD_PSEFL	Q51773 pseudomonas
314	6	3.4	1042	1	SYI_BORBU	O51773 borrelia bu	387	5	2.9	55	1	PHNS_DESVH	Q06173 desulfovibr
315	6	3.4	1043	1	P11D_MOUSE	Q35904 mus musculu	388	5	2.9	55	1	R322_STRCO	Q9rl50 streptomyce
316	6	3.4	1061	1	ECA1_ARATH	P92939 arabidopsis	389	5	2.9	56	1	YRZK_BACSU	O32040 bacillus su
317	6	3.4	1061	1	ECA4_ARATH	Q9xes1 arabidopsis	390	5	2.9	58	1	FER6_METJA	Q58041 methanococc
318	6	3.4	1089	1	DLP2_RAT	P97837 rattus norv	391	5	2.9	59	1	NAPE_PARP	Q56348 paracoccus
319	6	3.4	1146	1	YHC3_YEAST	P38742 saccharomyc	392	5	2.9	62	1	IT13_MOMCH	P09407 momordica c
320	6	3.4	1158	1	ALA1_ARATH	P98204 arabidopsis	393	5	2.9	63	1	IBB_VICFA	P24661 vicia faba
321	6	3.4	1163	1	TSC1_RAT	Q9xl36 rattus norv	394	5	2.9	64	1	RL28_CAMJE	Q9pi58 campylobact
322	6	3.4	1188	1	S3B1_SCHPO	Q10178 schizosacch	395	5	2.9	66	1	PEPA_SUNMU	P81497 suncus muri
323	6	3.4	1196	1	DNBI_HSV11	P04296 herpes simp	396	5	2.9	67	1	PN4A_PENVA	Q95nt0 penaeus van
324	6	3.4	1196	1	DNBI_HSV1F	P17469 herpes simp	397	5	2.9	67	1	PN4C_PENVA	Q963c3 penaeus van
325	6	3.4	1196	1	DNBI_HSV1K	P17470 herpes simp	398	5	2.9	67	1	YQ12_BACAN	Q9rn20 bacillus an

399	5	2.9	70	1	YBCJ_ECOLI	P45571	escherichia	472	5	2.9	111	1	CY32_DESDN	P38554	desulfovibr
400	5	2.9	71	1	HSTA_YEREN	P07593	yersinia en	473	5	2.9	112	1	CYC6_SYNEL	P56534	synechococc
401	5	2.9	73	1	Y43_BPR18	P18243	bacterioph	474	5	2.9	112	1	CYC6_SYNNU	Q9f119	synechococc
402	5	2.9	73	1	YPO4_NPVLD	P30326	lymantria d	475	5	2.9	112	1	HYP4_SYNTP6	P94160	synechococc
403	5	2.9	74	1	Y177_LEPIN	O50638	leptospira	476	5	2.9	112	1	KV1U_HUMAN	P01613	homo sapien
404	5	2.9	75	1	YML1_THIFE	P20087	thiobacillu	477	5	2.9	113	1	DAD1_CAEEL	P52872	caenorhabdi
405	5	2.9	76	1	UCRX_YEAST	P37299	saccharomyc	478	5	2.9	113	1	UMP2_ARATH	Q9zux4	arabidopsis
406	5	2.9	77	1	LEA2_BOVIN	Q95jtc3	bos taurus	479	5	2.9	114	1	YEUG_ECOLI	P33917	escherichia
407	5	2.9	77	1	LEA2_PIG	Q95jb4	sus scrofa	480	5	2.9	114	1	YFK5_YEAST	P43608	saccharomyc
408	5	2.9	77	1	RL28_NEIMA	Q9jgg3	neisseria m	481	5	2.9	114	1	YHIT_SYNTP7	P32084	synechococc
409	5	2.9	77	1	RL29_MYCBO	O06050	mycobacteri	482	5	2.9	114	1	YHIT_SYNTP3	P73481	synechocyst
410	5	2.9	77	1	RL29_MYCTU	P95057	mycobacteri	483	5	2.9	114	1	YJ76_SULSO	Q97wy4	sulfolobus
411	5	2.9	78	1	CINA_STRGV	P29827	streptover	484	5	2.9	115	1	DAD1_BETVE	Q9m3t9	betula verr
412	5	2.9	78	1	Y869_TREPA	O83839	treponema p	485	5	2.9	115	1	DAD1_CITUN	Q9zww7	citrus unsh
413	5	2.9	78	1	YUZF_BACSU	Q32097	bacillus su	486	5	2.9	115	1	RL19_TROW8	Q83i04	tropheryma
414	5	2.9	79	1	DINI_SERMA	Q9s380	serratia ma	487	5	2.9	115	1	RL19_TROWT	Q83g66	tropheryma
415	5	2.9	79	1	YJJ2_STRCO	O69880	streptomyce	488	5	2.9	116	1	DAD1_LYCES	Q9smc4	lycopersico
416	5	2.9	80	1	PERX_WHEAT	P15984	triticum ae	489	5	2.9	116	1	NUMM_RAT	P52504	rattus norv
417	5	2.9	80	1	RL29_MYCLE	Q32989	mycobacteri	490	5	2.9	116	1	REV_HV1A2	P04623	human immun
418	5	2.9	80	1	YK71_XYLFA	Q9pbr9	xylella fas	491	5	2.9	116	1	REV_HV1Y2	P35960	human immun
419	5	2.9	81	1	YP95_MYCTU	O50626	mycobacteri	492	5	2.9	116	1	RK21_MARPO	P06387	marchantia
420	5	2.9	82	1	COAB_BPPF1	P03621	bacterioph	493	5	2.9	116	1	RL17_SYNTP6	O24711	synechococc
421	5	2.9	83	1	COXJ_MOUSE	P48771	mus musculu	494	5	2.9	117	1	ELAF_HUMAN	P19957	homo sapien
422	5	2.9	83	1	COXJ_RAT	P35171	rattus norv	495	5	2.9	117	1	FRDD_YERPE	Q8zix8	yersinia pe
423	5	2.9	85	1	FDCS_HUMAN	Q8nfu4	homo sapien	496	5	2.9	118	1	RS13_BUCAI	P57569	buchnera ap
424	5	2.9	85	1	V12_BPT7	P03780	bacterioph	497	5	2.9	119	1	FRDD_ECOL6	Q8fal7	escherichia
425	5	2.9	86	1	YDCO_BACSU	P96632	bacillus su	498	5	2.9	119	1	FRDD_ECOLI	P03806	escherichia
426	5	2.9	87	1	CYC6_SYNLI	P00114	synechococc	499	5	2.9	119	1	FRDD_SALTY	Q8xfk9	salmonella
427	5	2.9	87	1	RS20_ZYMO	Q9z5v0	zymomonas m	500	5	2.9	120	1	MERD_SHIFL	P20102	shigella fl
428	5	2.9	87	1	VE4_HPVS1	P26548	human papil	501	5	2.9	120	1	VP3_CAV26	P54095	chicken ane
429	5	2.9	88	1	R31B_PASMU	Q9cp41	pasteurella	502	5	2.9	120	1	WNT9_ALOVU	P28107	alopias vul
430	5	2.9	88	1	YRCB_LACLA	Q9cf09	lactococcus	503	5	2.9	121	1	MERD_ACTICA	Q52110	acinetobact
431	5	2.9	89	1	NER_HAEIN	P46496	haemophilus	504	5	2.9	121	1	MERD_PSEAE	P06689	pseudomonas
432	5	2.9	89	1	Y4MD_RHISN	P55563	rhizobium s	505	5	2.9	121	1	MERD_SALTI	P94703	salmonella
433	5	2.9	91	1	YX57_MYCTU	O50386	mycobacteri	506	5	2.9	121	1	MERD_SERMA	P08654	serratia ma
434	5	2.9	92	1	MOTI_HORSE	O46617	equus cabal	507	5	2.9	121	1	VP3_CAV82	P54096	chicken ane
435	5	2.9	92	1	V12_BPT3	P07716	bacterioph	508	5	2.9	121	1	VP3_CAVC1	Q99152	chicken ane
436	5	2.9	95	1	YZX2_HUMAN	Q9y3y2	homo sapien	509	5	2.9	121	1	VP3_CAVCI	P54094	chicken ane
437	5	2.9	96	1	GATC_NEIMA	Q9jtz6	neisseria m	510	5	2.9	122	1	IBP6_BOVIN	Q05718	bos taurus
438	5	2.9	96	1	GATC_NEIMB	Q9jz00	neisseria m	511	5	2.9	122	1	SELH_HUMAN	Q8izq5	homo sapien
439	5	2.9	96	1	VREN_LAMBD	P03761	bacterioph	512	5	2.9	123	1	C59A_MOUSE	O55186	mus musculu
440	5	2.9	97	1	CSOA_THINE	P45689	thiobacillu	513	5	2.9	123	1	PSCA_MOUSE	P57096	mus musculu
441	5	2.9	98	1	CSOC_THINE	P45688	thiobacillu	514	5	2.9	123	1	YFJO_ECOLI	P52130	escherichia
442	5	2.9	99	1	ARG2_PHAU	P32292	phaseolus a	515	5	2.9	124	1	GTRA_SHIFL	P37785	shigella fl
443	5	2.9	99	1	GAS2_ARATH	P46688	arabidopsis	516	5	2.9	124	1	R35A_CAEEL	P49180	caenorhabdi
444	5	2.9	99	1	GAS3_ARATH	P46687	arabidopsis	517	5	2.9	124	1	RL22_GADMO	P52865	gadus morhu
445	5	2.9	99	1	IHFA_XYLFA	Q9pfd5	xylella fas	518	5	2.9	124	1	RL22_TREPA	Q83224	treponema p
446	5	2.9	99	1	IHFA_XYLFT	Q87ab7	xylella fas	519	5	2.9	124	1	YEEV_ECOLI	P76365	escherichia
447	5	2.9	99	1	PCP2_MOUSE	P12660	mus musculu	520	5	2.9	124	1	YG77_STRCO	Q9ad90	streptomyce
448	5	2.9	99	1	RL28_CAUCR	Q9aael	caulobacter	521	5	2.9	125	1	CYCP_RHOPA	P00149	rhodopseudo
449	5	2.9	99	1	YIIS_ECOLI	P32162	escherichia	522	5	2.9	126	1	C10_MOUSE	O35127	mus musculu
450	5	2.9	99	1	YQJK_ECOLI	Q47710	escherichia	523	5	2.9	126	1	U235_HUMAN	Q8wur7	homo sapien
451	5	2.9	102	1	RNPC_PENCH	P09647	penicillium	524	5	2.9	127	1	YC03_MYCPN	Q50284	mycoplasma
452	5	2.9	103	1	CCMK_PROMA	P96485	prochloroco	525	5	2.9	128	1	LSHB_STRCA	P80664	struthio ca
453	5	2.9	103	1	YD29_HAEIN	Q57255	haemophilus	526	5	2.9	128	1	RNPA_RHIME	Q92sf4	rhizobium m
454	5	2.9	104	1	YJA7_YEAST	P47080	saccharomyc	527	5	2.9	128	1	ZB14_MAIZE	P42856	zea mays (m
455	5	2.9	105	1	CORT_HUMAN	O00230	homo sapien	528	5	2.9	129	1	C59B_MOUSE	P58019	mus musculu
456	5	2.9	105	1	DBH_TREPA	O83278	treponema p	529	5	2.9	129	1	RPOZ_AGR75	Q8ugk8	agrobacteri
457	5	2.9	105	1	LE5A_GOSHI	P46521	gossypium h	530	5	2.9	129	1	RS9_BACST	P07842	bacillus st
458	5	2.9	105	1	LE5D_GOSHI	P46522	gossypium h	531	5	2.9	129	1	RS9_BACSU	P21470	bacillus su
459	5	2.9	105	1	NIGM_HUMAN	Q95178	homo sapien	532	5	2.9	129	1	RS9_WIGBR	Q8d362	wiggleswort
460	5	2.9	105	1	Y04A_METJA	P81302	methanococc	533	5	2.9	130	1	COXE_SCHPO	O74471	schizosacch
461	5	2.9	106	1	CYAY_YERPE	P46356	yersinia pe	534	5	2.9	130	1	RR11_CHIVU	P56359	chlorella v
462	5	2.9	106	1	FER6_RHOCA	P80306	rhodobacter	535	5	2.9	130	1	RS9_BACAA	Q81vp8	bacillus an
463	5	2.9	106	1	URE2_KLEAE	P18315	klebsiella	536	5	2.9	130	1	RS9_BACCR	Q81j12	bacillus ce
464	5	2.9	107	1	FLIE_RHILO	Q98hd8	rhizobium l	537	5	2.9	130	1	RS9_ENTFA	Q82z47	enterococcu
465	5	2.9	108	1	SDX_SULSO	Q97uv1	sulfolobus	538	5	2.9	130	1	RS9_LACLA	Q9cdg7	lactococcus
466	5	2.9	108	1	SDX_SULTO	Q9hh83	sulfolobus	539	5	2.9	130	1	RS9_LACPL	Q88xu7	lactobacill
467	5	2.9	110	1	CSOB_THINE	P45690	thiobacillu	540	5	2.9	130	1	RS9_STAAM	Q99s52	staphylococ
468	5	2.9	110	1	PT03_STYPL	P28195	styela plic	541	5	2.9	130	1	RS9_STAEP	Q8crj0	staphylococ
469	5	2.9	110	1	RL3E_METJA	P54061	methanococc	542	5	2.9	130	1	RS9_STR43	Q8e7e4	streptococc
470	5	2.9	110	1	SECG_ECOLI	P33582	escherichia	543	5	2.9	130	1	RS9_STR45	Q8ely6	streptococc
471	5	2.9	111	1	AQN1_PIG	P26322	sus scrofa	544	5	2.9	130	1	RS9_STRMU	Q8dw97	streptococc



545	130	1	RS9_STRPN	Q97sn4 streptococc	618	5	2.9	151	1	ZRAP_SALTI	Q8z331 salmonella
546	130	1	RS9_STRPY	Q99y08 streptococc	619	5	2.9	151	1	ZRAP_SALTY	Q919i0 salmonella
547	130	1	RS9_STRR6	Q8cwu4 streptococc	620	5	2.9	152	1	FTRC_MAIZE	P41347 zea mays (m
548	132	1	GCSH_MYCLE	O32920 mycobacteri	621	5	2.9	152	1	SP18_ARATH	O64644 arabidopsis
549	132	1	RNPA_MICLU	P21172 micrococcus	622	5	2.9	153	1	HS11_SOYBN	P02519 glycine max
550	132	1	VIF_HVILW	Q70623 human immun	623	5	2.9	153	1	HXC6_SHEEP	P49925 ovis aries
551	132	1	YOHJ_ECOLI	P33372 escherichia	624	5	2.9	153	1	MAUF_METFL	Q50418 methylobaci
552	134	1	ACPM_NEUCR	P11943 neurospora	625	5	2.9	153	1	RBM3_MOUSE	O89086 mus musculu
553	134	1	ATPE_MOOTH	O05434 moorella th	626	5	2.9	153	1	VE6_HPVS0	P36809 human papil
554	134	1	CAL2_RAT	P10093 rattus norv	627	5	2.9	153	1	VNSC_TPMV	Q9ws38 tupaia para
555	134	1	GCSH_MYCTU	Q50607 mycobacteri	628	5	2.9	153	1	VPG_BYDVP	P09513 barley yell
556	134	1	PRL4_HUMAN	Q16378 homo sapien	629	5	2.9	153	1	YMF8_ECOLI	P75965 escherichia
557	134	1	YF60_ARCFU	O28712 archaeoglob	630	5	2.9	154	1	AROQ_BUCBP	Q89ae0 buchnera ap
558	135	1	YGDB_ECOLI	P08370 escherichia	631	5	2.9	154	1	RISB_CHLCV	Q821p5 chlamydophi
559	136	1	RU VX_DEIRA	P09r12 deinococcus	632	5	2.9	154	1	SM20_SCHMA	P15845 schistosoma
560	137	1	CY2_RHOCA	P00094 rhodobacter	633	5	2.9	155	1	HOPD_SALTY	O68927 salmonella
561	137	1	MBB2_ECOLI	P07113 escherichia	634	5	2.9	155	1	PAND_AGRFS	Q8u9q9 agrobacteri
562	137	1	Y049_MYCTU	P71706 mycobacteri	635	5	2.9	155	1	PR11_PETCR	P19417 petroselinu
563	137	1	YH92_ARCFU	O28482 archaeoglob	636	5	2.9	155	1	PR13_PETCR	P19418 petroselinu
564	137	1	YZ58_PSEAE	Q9hy59 pseudomonas	637	5	2.9	155	1	RIFK_MOUSE	Q8cfv9 mus musculu
565	139	1	PSAD_ODOSI	P49481 odontella s	638	5	2.9	155	1	RL30_PYRAB	Q9v1v6 pyrococcus
566	139	1	PSAD_SKECO	O96800 skeletonema	639	5	2.9	155	1	RL30_PYRHO	O59440 pyrococcus
567	139	1	SPY1_HUMAN	O43609 homo sapien	640	5	2.9	155	1	Y4RL_RHLSN	P55645 rhizobium s
568	140	1	Y14K_CSMV	P18920 chloris str	641	5	2.9	156	1	LRP_RHIME	P56901 rhizobium m
569	140	1	Y228_METKA	Q8tyr4 methanopyru	642	5	2.9	156	1	VPG_PLRV	P10471 potato leaf
570	140	1	YUD1_STRAW	Q93hl9 streptomyce	643	5	2.9	156	1	VPG_PLRV1	P17524 potato leaf
571	141	1	CMAR_HUMAN	Q04762 homo sapien	644	5	2.9	156	1	VPG_PLRVVR	P17523 potato leaf
572	141	1	HBA_PHYCA	P09904 physeter ca	645	5	2.9	156	1	VPG_PLRVW	P11625 potato leaf
573	141	1	HBA_VULVU	P21200 vulpes vulp	646	5	2.9	157	1	GRPA_MAIZE	P10979 zea mays (m
574	141	1	YEDD_SALTY	Q06399 salmonella	647	5	2.9	157	1	GRP_DAUCA	Q03878 daucus caro
575	142	1	DUT_SPVKA	P32208 swinepox vi	648	5	2.9	157	1	RBM3_HUMAN	P98179 homo sapien
576	142	1	GRP1_SORBI	Q99069 sorghum bic	649	5	2.9	157	1	RISB_CHLMU	Q9plj4 chlamydia m
577	142	1	GTR2_BOVIN	P58351 bos taurus	650	5	2.9	157	1	RISB_CHLTR	O84737 chlamydia t
578	142	1	NXT2_HUMAN	Q9np18 homo sapien	651	5	2.9	157	1	VE6_HPVS2	P50776 human papil
579	142	1	PTP_ACIOJO	O52787 acinetobact	652	5	2.9	157	1	WH16_STRCO	P23157 streptomyce
580	142	1	V16K_BLRV	P19127 bean leafro	653	5	2.9	157	1	YE18_COXBU	Q83bt4 coxiella bu
581	142	1	YN8P_YEAST	P53737 saccharomyc	654	5	2.9	158	1	AB18_PEA	Q06930 pisum sativ
582	142	1	YQAN_BACSU	P45911 bacillus su	655	5	2.9	158	1	GRP2_ARATH	Q9svm8 arabidopsis
583	143	1	DSS4_YEAST	P32601 saccharomyc	656	5	2.9	158	1	LUXS_LACPL	Q88yi6 lactobacill
584	143	1	YK11_MYCTU	Q10846 mycobacteri	657	5	2.9	159	1	IF52_ARATH	Q93vp3 arabidopsis
585	144	1	IAAE_HORVU	P01096 hordeum vul	658	5	2.9	159	1	LSHB_MELGA	P45646 meleagris g
586	144	1	YBG7_PSEAE	O68560 pseudomonas	659	5	2.9	160	1	FRZA_MYXXA	P43498 myxococcus
587	145	1	SJ2B_HUMAN	P57105 homo sapien	660	5	2.9	160	1	PETD_LOTJA	Q9bbq5 lotus japon
588	146	1	CDD_MOUSE	P56389 mus musculu	661	5	2.9	160	1	RS7_RICPR	P41081 rickettsia
589	146	1	HBBC_CONCO	P83478 conger cong	662	5	2.9	160	1	U195_CABEL	O62252 caenorhabdi
590	146	1	RL26_BRARA	Q39411 brassica ra	663	5	2.9	161	1	RRAA_YERPE	Q8zjj7 yersinia pe
591	146	1	YDGK_ECOLI	P76180 escherichia	664	5	2.9	161	1	VAL2_CABEL	P34546 caenorhabdi
592	146	1	YPV9_METTF	P29586 methanobact	665	5	2.9	161	1	VATL_ASCSU	Q17046 ascaris suu
593	147	1	DUT_CHLCV	Q823q9 chlamydophi	666	5	2.9	162	1	CYPH_PARPR	O00845 paramecium
594	148	1	FER_ARATH	O04090 arabidopsis	667	5	2.9	162	1	RIFK_HUMAN	Q969g6 homo sapien
595	148	1	FER_RHOER	P16972 arabidopsis	668	5	2.9	162	1	RL24_HORVU	P50888 hordeum vul
596	148	1	LIMA_RHOER	Q9zag3 rhodococcus	669	5	2.9	162	1	VA49_VACCC	P21068 vaccinia vi
597	148	1	OLE1_PRUDU	Q43804 prunus dulc	670	5	2.9	162	1	VA49_VACCV	P31037 vaccinia vi
598	148	1	YLL2_EBV	P03199 epstein-bar	671	5	2.9	162	1	VA49_VARV	P33857 variola vir
599	148	1	YLL2_EBVA8	Q07285 epstein-bar	672	5	2.9	162	1	VENV_EAV	P28991 equine arte
600	148	1	YM25_YEAST	P40219 saccharomyc	673	5	2.9	162	1	YC87_RALSO	Q8xzxv8 ralstonia s
601	149	1	COAD_MYCPU	Q98rb3 mycoplasma	674	5	2.9	163	1	CIRP_XENLA	O93235 xenopus lae
602	149	1	H2A2_PEA	P40281 pisum sativ	675	5	2.9	163	1	NUIM_DROAI	P51926 drosophila
603	150	1	CRPD_CHLTR	O84449 chlamydia t	676	5	2.9	163	1	NUIM_DROAL	P51927 drosophila
604	150	1	H2A1_PEA	P25470 pisum sativ	677	5	2.9	163	1	NUIM_DROAM	P51928 drosophila
605	150	1	NAAA_BPT3	P20331 bacterioph	678	5	2.9	163	1	NUIM_DROBF	P51930 drosophila
606	150	1	NAAA_BPT7	P00806 bacterioph	679	5	2.9	163	1	NUIM_DROGU	P51931 drosophila
607	150	1	TAA1_VACCC	P20982 vaccinia vi	680	5	2.9	163	1	NUIM_DROMD	P51932 drosophila
608	150	1	TAA1_VACCV	P07610 vaccinia vi	681	5	2.9	163	1	NUIM_DROMI	P51933 drosophila
609	150	1	TAA1_VARV	P33814 variola vir	682	5	2.9	163	1	NUIM_DROSS	P51938 drosophila
610	150	1	YCCA_BACP3	P55815 bacillus ps	683	5	2.9	163	1	RL24_ARATH	P38666 arabidopsis
611	151	1	DKSA_BUCAP	Q8k9u5 buchnera ap	684	5	2.9	163	1	RU VX_CHLCV	Q822t1 chlamydophi
612	151	1	NGB_HUMAN	Q9np92 homo sapien	685	5	2.9	163	1	TPX1_STRPN	P72500 streptococc
613	151	1	SODC_BOVIN	P00442 bos taurus	686	5	2.9	164	1	FR1H_RABIT	P25915 oryctolagus
614	151	1	SODC_SHEEP	P09670 ovis aries	687	5	2.9	164	1	RL24_CICAR	O65743 cicer ariet
615	151	1	YFBV_ECOL6	Q8ffj2 escherichia	688	5	2.9	164	1	Y157_METAC	Q8tub6 methanosarc
616	151	1	YFBV_ECOLI	P77496 escherichia	689	5	2.9	165	1	COMD_METTH	O27274 methanobact
617	151	1	YFBV_SHIFL	Q83kb1 shigella fl	690	5	2.9	165	1	PDAL_METMA	P58889 methanosarc

691	1	RS5_CLOAB	165	2.9	5	Q97ej5	clostridium
692	1	DTD_RALSO	166	2.9	5	Q8y239	ralstonia s
693	1	GRPI_SINAL	166	2.9	5	P49310	sinapis alb
694	1	HBL1_ORYSA	166	2.9	5	O04986	oryza sativ
695	1	LSHB_COTJA	166	2.9	5	P45657	coturnix co
696	1	PAL_PSEPK	166	2.9	5	P43036	pseudomonas
697	1	RECX_SALTI	166	2.9	5	Q8z4d4	salmonella
698	1	RECX_SALTY	166	2.9	5	Q8zmk5	salmonella
699	1	VG04_BPP22	166	2.9	5	P26746	bacterioph
700	1	VSN1_NOCAR	166	2.9	5	P50186	nocardia ae
701	1	Y346_MYCPN	166	2.9	5	P47588	mycoplasma
702	1	Y516_SYNY3	166	2.9	5	P75257	mycoplasma
703	1	YL81_SCHPO	166	2.9	5	Q55837	synechocyst
704	1	COAD_DEIRA	167	2.9	5	O13995	schizosacch
705	1	HBL4_ORYSA	167	2.9	5	Q9xwm4	deinococcu
706	1	NADM_METJA	167	2.9	5	Q94ft7	oryza sativ
707	1	GRP2_SORBI	168	2.9	5	Q57961	methanococc
708	1	ISPF_RALSO	168	2.9	5	Q99070	sorghum bic
709	1	MOAC_DEIRA	168	2.9	5	Q8xyw2	ralstonia s
710	1	PTP_NPVAC	168	2.9	5	Q9rrc1	deinococcu
711	1	VAT_CERV	168	2.9	5	P24656	autographa
712	1	GR10_BRANA	169	2.9	5	P05397	carnation e
713	1	GRP2_SINAL	169	2.9	5	Q05966	brassica na
714	1	GRP8_ARATH	169	2.9	5	P49311	sinapis alb
715	1	HBL2_ORYSA	169	2.9	5	Q03251	arabidopsis
716	1	MINC_DEIRA	169	2.9	5	O04985	oryza sativ
717	1	PHAF_CYAPA	169	2.9	5	Q94ft8	oryza sativ
718	1	DEF1_BORBR	170	2.9	5	Q9rtk7	deinococcu
719	1	DEF1_BORPA	170	2.9	5	P48087	cyanophora
720	1	DEF2_BORPE	170	2.9	5	Q9a3g6	caulobacter
721	1	ING_SIGHI	170	2.9	5	Q7wqs9	bordetella
722	1	JDP_BOMMO	170	2.9	5	Q7wlv3	bordetella
723	1	JDP_MANSE	170	2.9	5	Q7vs88	bordetella
724	1	BB19_RABIT	171	2.9	5	Q9qxx2	sigmodon hi
725	1	HSCB_AZOV	171	2.9	5	Q9u6v7	bombyx mori
726	1	HSCB_ECOLI	171	2.9	5	Q9u6v6	manduca sex
727	1	HSCB_SALTY	171	2.9	5	Q05005	oryctolagus
728	1	HSCB_SHIFL	171	2.9	5	O69220	azotobacter
729	1	RECX_MYCLE	171	2.9	5	P36540	escherichia
730	1	YI62_METTH	171	2.9	5	Q8z4n1	salmonella
731	1	CIRP_HUMAN	171	2.9	5	Q8zn41	salmonella
732	1	CIRP_MOUSE	171	2.9	5	Q83qk3	shigella fl
733	1	RUVC_ECOLI	171	2.9	5	P37859	mycobacteri
734	1	GRAM_MOUSE	171	2.9	5	O27890	methanobact
735	1	HSCB_PSEPK	171	2.9	5	Q14011	homo sapien
740	1	NU6M_SCYCA	173	2.9	5	Q61413	mus musculu
741	1	NU6M_SQUAC	173	2.9	5	P24239	escherichia
742	1	RUVC_SALTY	173	2.9	5	P51437	mus musculu
743	1	SP12_YEAST	173	2.9	5	Q88pk5	pseudomonas
744	1	YA92_MYCPN	173	2.9	5	O79412	scyliorhinu
745	1	AROK_BUCBP	174	2.9	5	Q9zz43	squalus aca
746	1	RIMM_TREPA	174	2.9	5	Q8xgf7	salmonella
747	1	RS5_SYNEL	174	2.9	5	P17123	saccharomyc
748	1	Y4TN_RHISN	174	2.9	5	P75600	mycoplasma
749	1	YF48_ARCFU	174	2.9	5	P59488	buchnera ap
750	1	YY19_HUMAN	174	2.9	5	O83877	treponema p
751	1	EPO_CANFA	175	2.9	5	P59126	synechococc
752	1	THIM_WHEAT	175	2.9	5	P55668	rhizobium s
753	1	YJM2_YEAST	175	2.9	5	O28724	archaeoglob
754	1	YMF3_CAEEL	175	2.9	5	P09002	homo sapien
755	1	GRP7_ARATH	176	2.9	5	P33707	canis famil
756	1	TX13_MOUSE	177	2.9	5	Q9z2p21	tritricum ae
757	1	VC08_VACCV	177	2.9	5	P47019	saccharomyc
758	1	VEG2_RAT	177	2.9	5	P34463	caenorhabdi
759	1	YH84_ARCFU	177	2.9	5	Q03250	arabidopsis
760	1	AAC2_PROST	178	2.9	5	O54841	mus musculu
761	1	FXY5_HUMAN	178	2.9	5	P17364	vaccinia vi
762	1	NADM_METTH	178	2.9	5	P41244	rattus norv
763	1	RL1X_ARATH	178	2.9	5	Q28490	archaeoglob

178	1	RL1Y_ARATH	178	2.9	5	764	
178	1	SSB2_ECOLI	178	2.9	5	765	
178	1	TRPG_ARCFU	178	2.9	5	766	
178	1	YGYS_YEAST	178	2.9	5	767	
179	1	14P_BOVIN	179	2.9	5	768	
179	1	E320_ADE03	179	2.9	5	769	
179	1	HTRC_ECOLI	179	2.9	5	770	
179	1	IF3_LEPIN	179	2.9	5	771	
179	1	RK5_ASILLO	179	2.9	5	772	
179	1	STC_ONCKE	179	2.9	5	773	
180	1	FRIH_BOVIN	180	2.9	5	774	
180	1	PYRE_METH	180	2.9	5	775	
180	1	YW60_XANCP	180	2.9	5	776	
180	1	YY06_XANAC	180	2.9	5	777	
181	1	APH1_SCHPO	181	2.9	5	778	
181	1	FRIH_MOUSE	181	2.9	5	779	
181	1	FRIH_RAT	181	2.9	5	780	
181	1	IGF2_HORSE	181	2.9	5	781	
181	1	RRP3_HORVU	181	2.9	5	782	
181	1	Y4WG_RHISN	181	2.9	5	783	
182	1	PDAD_THEAC	182	2.9	5	784	
182	1	PDAD_THEVO	182	2.9	5	785	
182	1	PGRP_TRINI	182	2.9	5	786	
182	1	VG37_BPMU	182	2.9	5	787	
183	1	AAC1_DICDI	183	2.9	5	788	
183	1	ORN_COXBU	183	2.9	5	789	
183	1	RS10_ORYSA	183	2.9	5	790	
184	1	HRPL_PSESY	184	2.9	5	791	
184	1	SPC3_YEAST	184	2.9	5	792	
184	1	VC08_VACCC	184	2.9	5	793	
184	1	Y583_METH	184	2.9	5	794	
184	1	YFEE_YERPE	184	2.9	5	795	
185	1	ADH_ANAPL	185	2.9	5	796	
185	1	FRIH_CRIGR	185	2.9	5	797	
185	1	NP77_XENLA	185	2.9	5	798	
185	1	PAPA_ECOLI	185	2.9	5	799	
185	1	PSAF_PORPU	185	2.9	5	800	
185	1	YSD2_CAEEL	185	2.9	5	801	
186	1	GRPE_BACSU	186	2.9	5	802	
186	1	NIAM_BOVIN	186	2.9	5	803	
186	1	NIAM_HUMAN	186	2.9	5	804	
187	1	LPPJ_MYCTU	187	2.9	5	805	
187	1	PYRE_METAC	187	2.9	5	806	
187	1	PYRE_METMA	187	2.9	5	807	
187	1	RL5_MYCLE	187	2.9	5	808	
187	1	RL5_MYCTU	187	2.9	5	809	
187	1	YG18_ARCFU	187	2.9	5	810	
188	1	APT_HALN1	188	2.9	5	811	
188	1	NRFK_PASMU	188	2.9	5	812	
188	1	PIGH_HUMAN	188	2.9	5	813	
189	1	HPPA_AGRTU	189	2.9	5	814	
189	1	Y415_XYLFT	189	2.9	5	815	
189	1	YB24_XYLFA	189	2.9	5	816	
189	1	YI35_PASMU	189	2.9	5	817	
190	1	COAT_WCMVM	190	2.9	5	818	
190	1	DSR6_HUMAN	190	2.9	5	819	
190	1	MOBA_RHOCA	190	2.9	5	820	
190	1	PP28_HCMVA	190	2.9	5	821	
190	1	PPIA_ECOLI	190	2.9	5	822	
190	1	PPIA_ERWCH	190	2.9	5	823	
190	1	PPIA_SALTY	190	2.9	5	824	
190	1	RNH2_SYNY3	190	2.9	5	825	
191	1	TFAQ_ECOLI	191	2.9	5	826	
192	1	AMEY_HUMAN	192	2.9	5	827	
192	1	COMK_BACSU	192	2.9	5	828	
192	1	EPO_FELCA	192	2.9	5	829	
192	1	PYRR_CORGL	192	2.9	5	830	
192	1	RM11_HUMAN	192	2.9	5	831	
192	1	RS4B_CLOAB	192	2.9	5	832	
192	1	VIF_HV1B1	192	2.9	5	833	
192	1	VIF_HV1P5	192	2.9	5	834	
192	1	VIF_HV1NA	192	2.9	5	835	
192	1	Y4GC_RHISN	192	2.9	5	836	

Q9lud4	arabidopsis
P28043	escherichia
O28670	archaeoglob
P53071	saccharomyc
O18883	bos taurus
P11321	human adeno
P27375	escherichia
Q8f6q9	leptospira
P14757	astasia lon
P43647	oncorhynch
O46414	bos taurus
O27888	methanobact
Q8p5s6	xanthomonas
Q8ph54	xanthomonas
P49776	schizosacch
P09528	mus muscull
P19132	rattus norv
P51459	equus cabal
O48609	hordeum vul
P55685	rhizobium s
Q9hk30	thermoplasm
Q97an7	thermoplasm
O76537	trichoplusi
Q9t1v8	bacterioph
P14195	dictyosteli
Q83c93	coxiella bu
Q9ayp4	oryza sativ
P37929	pseudomonas
Q12133	saccharomyc
P21041	vaccinia vi
O26683	methanobact
Q56956	yersinia pe
P30350	anas platyr
P29389	cricketulus
P55869	xenopus lae
P04127	escherichia
P51193	porphyra pu
Q10121	caenorhabdi
P15874	bacillus su
Q02372	bos taurus
O95169	homo sapien
Q10688	mycobacteri
P58859	methanosarc
Q8q0j4	methanosarc
O32995	mycobacteri
P95064	mycobacteri
O28655	archaeoglob
Q9hrt1	halobacteri
Q9cpk9	pasteurella
Q14442	homo sapien
Q8vpz0	agrobacteri
Q87ea3	xylella fas
Q9pea4	xylella fas
Q9ck03	pasteurella
P09502	white clove
P57055	homo sapien
Q9x7k0	rhodobacter
P13200	human cytom
P20752	escherichia
O53021	erwinia chr
P20753	salmonella
P72657	synechocyst
P76155	escherichia
Q99218	homo sapien
P40396	bacillus su
P33708	felis silve
P59011	corynebacte
Q9y3b7	homo sapien
Q97j08	clostridium
P03401	human immun
P04598	human immun
P31820	human immun
P55459	rhizobium s



837	5	2.9	193	1	SODF_BACPR	P53638 bacteroides	910	5	2.9	205	1	RUVA_AGR75	Q8u9k5 agrobacteri
838	5	2.9	193	1	YG89_OCEIH	P59250 oceanobacil	911	5	2.9	205	1	RUVA_BRUME	Q8yiv6 brucella me
839	5	2.9	193	1	YK02_YEAST	P36042 saccharomyc	912	5	2.9	205	1	SC11_OMMSL	P18426 ommastrephe
840	5	2.9	193	1	YNES_BACSU	Q45064 bacillus su	913	5	2.9	206	1	SJ2B_RAT	Q9wvj4 rattus norv
841	5	2.9	194	1	CT45_HUMAN	Q9y3b1 homo sapien	914	5	2.9	207	1	AHYI_AERHY	Q44058 aeromonas h
842	5	2.9	194	1	PABA_BACSU	P28819 b para-amin	915	5	2.9	207	1	ASAI_AERSA	P70774 aeromonas s
843	5	2.9	194	1	PDX1_DROME	Q9v3p0 drosophila	916	5	2.9	207	1	CYSC_VIBVU	Q8de75 vibrio vuln
844	5	2.9	194	1	PSPC_RAT	P11685 rattus norv	917	5	2.9	207	1	RL13_CAEEL	P91128 caenorhabdi
845	5	2.9	194	1	TFA_LAMBD	P03740 bacterioph	918	5	2.9	208	1	COAT_WCMVO	P15406 white clove
846	5	2.9	194	1	YS67_DEIRA	Q9ryy6 deinococcus	919	5	2.9	208	1	ENGB_HELPJ	Q9zj42 helicobacte
847	5	2.9	195	1	NRTN_MOUSE	P97463 mus musculu	920	5	2.9	208	1	ENGB_HELPY	O26087 helicobacte
848	5	2.9	195	1	ROA1_BOVIN	P09867 bos taurus	921	5	2.9	208	1	IL6_FELCA	P41683 felis silve
849	5	2.9	195	1	YZG1_CAEEL	P55326 caenorhabdi	922	5	2.9	208	1	PAFR_MACMU	P35366 macaca mula
850	5	2.9	196	1	GCH2_ECOL6	Q8fhu5 escherichia	923	5	2.9	208	1	RS4_THETH	P80373 thermus the
851	5	2.9	196	1	GCH2_ECOLI	P25523 escherichia	924	5	2.9	208	1	TRPF_NEIMA	Q9jvd1 neisseria m
852	5	2.9	196	1	GCH2_SALTY	Q8xfy7 salmonella	925	5	2.9	208	1	TRPF_NEIMB	Q9k0c6 neisseria m
853	5	2.9	196	1	GCH2_YERPE	Q8zef0 yersinia pe	926	5	2.9	208	1	UPP_NEIMA	Q9jv58 neisseria m
854	5	2.9	196	1	HADB_BURPI	Q53009 burkholderi	927	5	2.9	208	1	UPP_NEIMB	Q9k048 neisseria m
855	5	2.9	196	1	NODC_RHISM	P04679 rhizobium s	928	5	2.9	208	1	YE68_MYCPN	P75315 mycoplasma
856	5	2.9	196	1	PGRP_BOWMO	Q9xtn0 bombyx mori	929	5	2.9	209	1	CHEC_BACSU	P40403 bacillus su
857	5	2.9	196	1	R18A_HUMAN	Q9nvs2 homo sapien	930	5	2.9	209	1	DEF2_COXBU	Q83ak6 coxiella bu
858	5	2.9	196	1	R18A_MOUSE	Q99n85 mus musculu	931	5	2.9	209	1	EFA2_MOUSE	P52801 mus musculu
859	5	2.9	196	1	RR4_RHAHU	P36470 rhapis humi	932	5	2.9	209	1	HXA7_XENLA	P09071 xenopus lae
860	5	2.9	196	1	Y927_TREPA	O83897 treponema p	933	5	2.9	209	1	UPP_LACPL	Q9re01 lactobacill
861	5	2.9	196	1	YK59_ARCFU	O28220 archaeoglob	934	5	2.9	209	1	UPP_STRAA3	Q8e3w9 streptococc
862	5	2.9	197	1	GCH2_NEIMA	Q9ju99 neisseria m	935	5	2.9	209	1	CTG2_HUMAN	O8dst6 streptococc
863	5	2.9	197	1	GCH2_NEIMB	Q9jz78 neisseria m	936	5	2.9	210	1	CTG2_HUMAN	O75638 homo sapien
864	5	2.9	197	1	GCH2_WIGBR	Q8d2j0 wiggleswort	937	5	2.9	210	1	YH90_SYNY3	P72793 synechocyst
865	5	2.9	197	1	T4S5_HUMAN	O14894 homo sapien	938	5	2.9	210	1	YQS3_CAEEL	Q09311 caenorhabdi
866	5	2.9	198	1	JDP1_MOUSE	Q9r022 mus musculu	939	5	2.9	211	1	END3_TREPA	O83754 treponema p
867	5	2.9	198	1	LMBV_CHICK	Q01636 gallus gall	940	5	2.9	211	1	GLP1_SINAL	P45854 sinapis alb
868	5	2.9	198	1	Y0A4_BACCR	Q812z6 bacillus ce	941	5	2.9	211	1	GPMA_RHIME	Q92t25 rhizobium m
869	5	2.9	198	1	YC84_LISMO	Q8y7j3 listeria mo	942	5	2.9	211	1	LEXA_XYLFA	Q9ph24 xylella fas
870	5	2.9	198	1	YD23_LISIN	Q92C68 listeria in	943	5	2.9	211	1	LEXA_XYLFT	Q87f45 xylella fas
871	5	2.9	198	1	YM47_YEAST	Q03233 saccharomyc	944	5	2.9	211	1	RL4_LEPIN	Q9xd35 leptospira
872	5	2.9	199	1	CN06_HUMAN	Q8taal1 homo sapien	945	5	2.9	211	1	UPP_LACLA	Q9cec9 lactococcus
873	5	2.9	199	1	COBS_HALN1	Q9hpl2 halobacteri	946	5	2.9	211	1	UPP_LACLC	P50926 lactococcus
874	5	2.9	199	1	GCH2_BUCBP	P59555 buchnera ap	947	5	2.9	211	1	V33P_ADE40	P11805 human adeno
875	5	2.9	199	1	HMG4_HUMAN	O15347 homo sapien	948	5	2.9	211	1	YA03_ARCFU	O29259 archaeoglob
876	5	2.9	200	1	DSBA_VIBCH	P32557 vibrio chol	949	5	2.9	212	1	ATP6_TROMO	O03359 tropidurus
877	5	2.9	200	1	NADD_SYNY3	P73246 synechocyst	950	5	2.9	212	1	BIOD_AGR75	Q8u8t9 agrobacteri
878	5	2.9	200	1	NUPL_XENLA	P05221 xenopus lae	951	5	2.9	212	1	NUIM_CAEEL	Q22619 caenorhabdi
879	5	2.9	201	1	NT5C_HUMAN	Q8tcd5 homo sapien	952	5	2.9	212	1	NUSB_SYNEL	Q8dks0 synechococc
880	5	2.9	201	1	RHLI_PSEAE	P54291 pseudomonas	953	5	2.9	212	1	RECX_CLOPE	O8xj38 clostridium
881	5	2.9	201	1	RUVA_DEIRA	Q9ruv7 deinococcus	954	5	2.9	212	1	RPCY_YEAST	P35718 saccharomyc
882	5	2.9	201	1	SDC2_HUMAN	P34741 homo sapien	955	5	2.9	212	1	YC25_ARCFU	O29043 archaeoglob
883	5	2.9	202	1	AMEL_MONDO	Q28462 monodelphis	956	5	2.9	212	1	YN87_YEAST	P53719 saccharomyc
884	5	2.9	202	1	CTD6_HUMAN	Q9byl1 homo sapien	957	5	2.9	213	1	EFA2_HUMAN	O43921 homo sapien
885	5	2.9	202	1	LIF_RAT	P17777 rattus norv	958	5	2.9	213	1	GTH1_MAIZE	P12653 zea mays (m
886	5	2.9	202	1	RS4_SYNEL	P59134 synechococc	959	5	2.9	213	1	HIS1_BACSU	O34520 bacillus su
887	5	2.9	202	1	VIC2_AGRRH	P13460 agrobacteri	960	5	2.9	213	1	ILL1_HUMAN	P15814 homo sapien
888	5	2.9	202	1	VIC2_AGR75	P07166 agrobacteri	961	5	2.9	213	1	LEXA_XANAC	O86050 xanthomonas
889	5	2.9	202	1	VIC2_AGR7U	P06666 agrobacteri	962	5	2.9	213	1	LEXA_XANCA	Q93mq9 xanthomonas
890	5	2.9	203	1	EM24_YEAST	P32803 saccharomyc	963	5	2.9	213	1	LEXA_XANCP	Q8p9x2 xanthomonas
891	5	2.9	203	1	ET1_FIG	P09558 sus scrofa	964	5	2.9	213	1	MDCG_XANCP	Q8p4u3 xanthomonas
892	5	2.9	203	1	GCH2_SHEON	Q8edd1 shewanella	965	5	2.9	213	1	PSQ2_MAIZE	Q41806 zea mays (m
893	5	2.9	203	1	GT_PROMI	P15214 proteus mir	966	5	2.9	213	1	TERX_SERMA	P75012 serratia ma
894	5	2.9	203	1	KTHY_CHLTR	O84191 chlamydia t	967	5	2.9	213	1	YM06_YEAST	Q04477 saccharomyc
895	5	2.9	203	1	LIF_MOUSE	P09056 mus musculu	968	5	2.9	214	1	ACUB_BACSU	P39066 bacillus su
896	5	2.9	203	1	RS3A_METAC	Q8tki9 methanosarc	969	5	2.9	214	1	ARA4_ARATH	P28187 arabidopsis
897	5	2.9	204	1	COAE_VIBPA	Q87lt4 vibrio para	970	5	2.9	214	1	MPP9_HUMAN	Q99550 homo sapien
898	5	2.9	204	1	LIPB_NITEU	Q82uj6 nitrosomona	971	5	2.9	214	1	MSA2_SYNY3	P72800 synechocyst
899	5	2.9	204	1	RS4_STRCO	Q9kxp5 streptomyce	972	5	2.9	214	1	Y540_AQUAE	O66819 aquifex aeo
900	5	2.9	204	1	T413_HUMAN	O95857 homo sapien	973	5	2.9	215	1	LB31_ARATH	O81322 arabidopsis
901	5	2.9	204	1	YK13_YEAST	P36079 saccharomyc	974	5	2.9	215	1	LEUD_XANAC	Q8ph04 xanthomonas
902	5	2.9	204	1	YM09_PARTE	P15610 paramecium	975	5	2.9	215	1	LEUD_XANCP	Q8p5k9 xanthomonas
903	5	2.9	205	1	CYSC_VIBPA	Q87sx6 vibrio para	976	5	2.9	215	1	NODB_RHILT	P04676 rhizobium l
904	5	2.9	205	1	DMT1_TRASC	P57690 trachemys s	977	5	2.9	215	1	RNI_AERHY	Q07465 aeromonas h
905	5	2.9	205	1	GCH2_PSEAE	Q9hwyl pseudomonas	978	5	2.9	216	1	DEF1_STRAW	Q82iv0 streptomyce
906	5	2.9	205	1	GCH2_PSEPK	Q88qhl pseudomonas	979	5	2.9	216	1	DEF4_STRCO	Q9k4a0 streptomyce
907	5	2.9	205	1	GCH2_PSESM	Q889q3 pseudomonas	980	5	2.9	216	1	FGFJ_HUMAN	Q95750 homo sapien
908	5	2.9	205	1	NOX_THETH	Q60049 thermus the	981	5	2.9	216	1	NODB_RHILV	P04339 rhizobium l
909	5	2.9	205	1	RNS9_HUMAN	P60153 homo sapien	982	5	2.9	216	1	O1E5_HUMAN	Q9um60 homo sapien

983	5	2.9	216	1	01E6_HUMAN	O43884 homo sapien
984	5	2.9	216	1	01FA_HUMAN	O43882 homo sapien
985	5	2.9	216	1	RR3_GUITH	O46900 guillardia
986	5	2.9	216	1	YJ87_ARCFU	O28292 archaeoglob
987	5	2.9	216	1	YS86_CABEL	Q09382 caenorhabdi
988	5	2.9	217	1	CREB_HYDAT	P51985 hydra atten
989	5	2.9	217	1	DEF1_BIFLO	Q8G534 bifidobacte
990	5	2.9	217	1	HXB7_MOUSE	P09024 mus musculu
991	5	2.9	217	1	NHAB_PSEPU	P97052 pseudomonas
992	5	2.9	217	1	NODB_RHIME	P02963 rhizobium m
993	5	2.9	217	1	PSQ1_MAIZE	Q41048 zea mays (m
994	5	2.9	217	1	RNH2_AGR75	Q8uhg2 agrobacteri
995	5	2.9	217	1	SOM2_MACMU	Q07370 macaca mula
996	5	2.9	217	1	TYSY_METAC	Q8thh4 methanosarc
997	5	2.9	217	1	TYSY_METMA	Q8pxi1 methanosarc
998	5	2.9	217	1	V33P_ADE41	P19416 human adeno
999	5	2.9	217	1	VTI1_YEAST	Q04338 saccharomyc
1000	5	2.9	217	1	YD32_MYCLE	P53424 mycobacteri

ALIGNMENTS

RESULT 1  
PGPL HUMAN  
ID PGPL HUMAN STANDARD; PRT; 576 AA.  
AC Q96PD5; Q96N74;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE N-acetylmuramoyl-L-alanine amidase precursor (EC 3.5.1.28)  
DE (peptidoglycan recognition protein long) (PGRP-L) (UNQ3103/PRO10102).  
DE PGRPL OR PGLYRPL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE DISTRIBUTION.  
MEDLINE=21433985; PubMed=11461926;  
Liu C., Xu Z., Gupta D., Dziarski R.;  
"Peptidoglycan recognition proteins: a novel family of four human  
innate immunity pattern recognition molecules."  
J. Biol. Chem. 276:34686-34694(2001).  
[2]  
SEQUENCE FROM N.A. (ISOFORM 2).  
TISSUE=Liver;  
Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,  
Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,  
Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,  
Masuho Y., Nagai K., Isogai T.;  
"NEDO human cDNA sequencing project."  
Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A. (ISOFORM 1).  
MEDLINE=22887296; PubMed=12975309;  
Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,  
Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
Godowski P.;  
"The secreted protein discovery initiative (SPDI), a large-scale  
effort to identify novel human secreted and transmembrane proteins: a  
bioinformatics assessment";  
Genome Res. 13:2265-2270(2003).

RNA  
RP SEQUENCE OF 22-36.  
RX MEDLINE=95392215; PubMed=7663175;  
RA De Pauw P., Neyt C., Vanderwinkel E., Wattiez R., Falmagne P.;  
RT "Characterization of human serum N-acetylmuramyl-L-alanine amidase  
purified by affinity chromatography."  
RL Protein Expr. Purif. 6:371-378(1995).  
[5]  
RN FUNCTION, AND MUTAGENESIS.  
RP PubMed=14506276;  
RX Wang Z.-M., Li X., Cocklin R.R., Wang M., Wang M., Fukase K.,  
RA Inamura S., Kusumoto S., Gupta D., Dziarski R.;  
RT "Human peptidoglycan recognition protein-L (PGRP-L) is an N-  
acetylmuramoyl-L-alanine amidase."  
RL J. Biol. Chem. 278:0-0(2003).  
[6]  
RN CARBOHYDRATE-LINKAGE SITE ASN-485.  
RP MEDLINE=22660472; PubMed=12754519;  
RX Zhang H., Li X.-J., Martin D.B., Aebbersold R.;  
RT "Identification and quantification of N-linked glycoproteins using  
hydrazide chemistry, stable isotope labeling and mass spectrometry."  
RL Nat. Biotechnol. 21:660-666(2003).  
CC -!- FUNCTION: May plays a scavenger role by digesting biologically  
CC active peptidoglycan (PGN) into biologically inactive fragments.  
CC Has no direct bacteriolytic activity.  
CC -!- CATALYTIC ACTIVITY: Hydrolyzes the link between N-acetylmuramoyl  
CC residues and L-amino acid residues in certain bacterial cell-wall  
CC glycopeptides.  
CC -!- COFACTOR: Zinc (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted and membrane-associated.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q96PD5-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q96PD5-2; Sequence=VSP\_008964;  
CC Note=May be due to an intron retention. No experimental  
CC confirmation available;  
CC -!- TISSUE SPECIFICITY: Strongly expressed in liver and fetal liver.  
CC Expressed to a much lesser extent in trverse colon, lymph nodes,  
CC heart, thymus, pancreas, descending colon, stomach and testis.  
CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase  
CC family 2.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AF384856; AAL05629.1; -.  
CC EMBL; AK055882; BAB71034.1; -.  
CC EMBL; AY358156; AAQ88523.1; -.  
CC MIM; 608199; -.  
CC InterPro; IPR002502; Amidase\_2.  
CC InterPro; IPR006619; PGRP.  
CC Pfam; PF01510; Amidase\_2; 1.  
CC SMART; SM00644; Ami\_2; 1.  
CC SMART; SM00701; PGRP; 1.  
CC Hydrolase; Immune response; Metal-binding; Zinc; Signal; Glycoprotein;  
CC Alternative splicing.  
CC SIGNAL 1 21  
CC CHAIN 22 576  
CC METAL 411 411  
CC METAL 447 447  
CC METAL 522 522  
CC METAL 530 530  
CC DISULFID 419 425  
CC CARBOHYD 77 77  
CC CARBOHYD 367 367  
CC CARBOHYD 485 485  
-----  
CC N-ACETYLMURAMOYL-L-ALANINE AMIDASE.  
CC ZINC (BY SIMILARITY).  
CC ZINC (BY SIMILARITY).  
CC ZINC (BY SIMILARITY).  
CC ZINC (BY SIMILARITY).  
CC BY SIMILARITY.  
CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC N-LINKED (GLCNAC. . .).



```
FT VARSPLIC 548 576 TVKRPARSVSKRSRREPPRTLPATDLO -> VSLRSLHY
FT TARRPSVTSSSTRELPPACNSCARTASARPPTSRRHHVSGN
FT LGPAFAGHSAGNIPDPVTSAYAAQAQTPQACFPFSS
FT (in isoform 2).
FT /FTid=VSP 008964.
FT H->A: NO EFFECT ON AMIDASE ACTIVITY.
FT C->A: ABOLISHES AMIDASE ACTIVITY.
FT MUTAGEN 411 411
FT MUTAGEN 419 419
FT MUTAGEN 436 436
FT MUTAGEN 442 442
FT MUTAGEN 447 447
FT MUTAGEN 530 530
FT MUTAGEN 530 530
FT CONFLICT 46 46
FT CONFLICT 99 99
FT CONFLICT 270 270
FT CONFLICT 394 394
FT CONFLICT 448 448
FT SEQUENCE 576 AA; 62216 MW; 73EA8713DC54F85A CRC64;
Query Match 50.0%; Score 87; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 8e-84;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGHWVGAAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVDTLPSCAVRAGLLRPDYALL 60
Db 461 RGHWVGAAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVDTLPSCAVRAGLLRPDYALL 520
QY 61 GHRQLVRTDCPGDALFDLLRTWPHFTA 87
Db 521 GHRQLVRTDCPGDALFDLLRTWPHFTA 547
RESULT 2
PGPL_PIG STANDARD; PRT; 598 AA.
AC Q866Y3; Q866Y4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE N-acetylmuramoyl-L-alanine amidase precursor (EC 3.5.1.28)
DE (peptidoglycan recognition protein long) (PGRP-L).
GN PGRPL.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RA Sang Y., Roess C.R., Blecha F.;
RT "Characterization of porcine peptidoglycan recognition proteins: gene
RT cloning and regulation on innate immunity.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May plays a scavenger role by digesting biologically
CC active peptidoglycan (PGN) into biologically inactive fragments.
CC Has no direct bacteriolytic activity.
CC -!- CATALYTIC ACTIVITY: Hydrolyzes the link between N-acetylmuramoyl
CC residues and L-amino acid residues in certain bacterial cell-wall
CC glycopeptides.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted and membrane-associated.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=B;
CC IsoId=Q866Y3-1; Sequence=Displayed;
CC Name=A;
CC IsoId=Q866Y3-2; Sequence=VSP_009082, VSP_009083;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase
CC family 2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF541955; AA041115.1; -.
DR EMBL; AF541956; AA041116.1; -.
DR InterPro; IPR002502; Amidase_2.
DR InterPro; IPR006619; PGRP.
DR Pfam; PF01510; Amidase_2; 1.
DR SMART; SM00644; Ami_2; 1.
DR SMART; SM00701; PGRP; 1.
KW Hydrolase; Immune response; Metal-binding; Zinc; Signal; Glycoprotein;
KW Alternative splicing.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 598 N-ACETYLMURAMOYL-L-ALANINE AMIDASE.
FT METAL 433 433 ZINC (BY SIMILARITY).
FT METAL 469 469 ZINC (BY SIMILARITY).
FT METAL 544 544 ZINC (BY SIMILARITY).
FT METAL 552 552 ZINC (BY SIMILARITY).
FT DISULFID 441 447 BY SIMILARITY.
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 507 507 N-LINKED (GLCNAC. .) (BY SIMILARITY).
FT VARSPLIC 1 346 Missing (in isoform A).
FT VARSPLIC 347 356 /FTid=VSP_009082.
FT VARSPLIC 347 356 /FTid=VSP_009083.
FT SEQUENCE 598 AA; 64593 MW; FCD237A9F105DDB CRC64;
Query Match 12.6%; Score 22; DB 1; Length 598;
Best Local Similarity 100.0%; Pred. No. 3.9e-15;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 PDYALLGHRQLVRTDCPGDALF 76
Db 537 PDYALLGHRQLVRTDCPGDALF 558
RESULT 3
PGPL_MOUSE STANDARD; PRT; 530 AA.
AC Q8VCS0; Q8K4I8; Q9QXZ1; Q9QXZ2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE N-acetylmuramoyl-L-alanine amidase precursor (EC 3.5.1.28)
DE (peptidoglycan recognition protein long) (PGRP-L) (tagL).
GN PGRPL OR PGLYRPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 23-37.
RC STRAIN=C57BL/6J;
RX MEDLINE=22705304; PubMed=12821140;
RA Gelius E., Persson C., Karlsson J., Steiner H.;
RT "A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-
RT L-alanine amidase activity.";
RL Biochem. Biophys. Res. Commun. 306:988-994(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).
RX MEDLINE=22447357; PubMed=12559944;
RA Kibardin A.V., Mirkina I.I., Baranova E.V., Zakeyeva I.R.,
RA Georgiev G.P., Kiselev S.L.;
RT "The differentially spliced mouse tagL gene, homolog of tag7/PGRP
RT gene family in mammals and Drosophila, can recognize Gram-positive
RT and Gram-negative bacterial cell wall independently of T phage
RT lysozyme homology domain.";
RL J. Mol. Biol. 326:467-474(2003).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -!- FUNCTION: May plays a scavenger role by digesting biologically  
CC active peptidoglycan (PGN) into biologically inactive fragments.  
CC  
CC -!- Has no direct bacteriolytic activity.  
CC -!- CATALYTIC ACTIVITY: Hydrolyzes the link between N-acetylmuramoyl  
CC residues and L-amino acid residues in certain bacterial cell-wall  
CC glycopeptides.  
CC -!- COFACTOR: Zinc (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted and membrane-associated.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=TagL-alpha;  
CC IsoId=Q8VCS0-1; Sequence=Displayed;  
CC Name=2; Synonyms=TagL-beta;  
CC IsoId=Q8VCS0-2; Sequence=VSP\_009081;  
CC Name=3; Synonyms=TagL-epsilon;  
CC IsoId=Q8VCS0-3; Sequence=VSP\_009079, VSP\_009080;  
CC -!- TISSUE SPECIFICITY: Strongly expressed in liver and fetal liver.  
CC -!- SIMILARITY: Belongs to the N-acetylmuramoyl-L-alanine amidase  
CC family 2.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AY282722; AAP22283.1; -.  
CC EMBL; AF392055; AAM73674.1; -.  
CC EMBL; AF149837; AAF22233.1; -.  
CC EMBL; AF149838; AAF22234.1; -.  
CC EMBL; BC019396; AAH19396.1; -.  
CC MGD; MGI:1928099; Pglyrpl.  
CC InterPro; IPR002502; Amidase\_2.  
CC InterPro; IPR006619; PGRP.  
CC Pfam; PF01510; Amidase\_2; 1.  
CC SMART; SM00644; Ami\_2; 1.  
CC SMART; SM00701; PGRP; 1.  
KW Hydrolase; Immune response; Metal-binding; Zinc; Signal; Glycoprotein;  
KW Alternative splicing.  
FT SIGNAL 1 22 N-ACETYLMURAMOYL-L-ALANINE AMIDASE.  
FT CHAIN 23 530 ZINC (BY SIMILARITY).  
FT METAL 391 391 ZINC (BY SIMILARITY).  
FT METAL 427 427 ZINC (BY SIMILARITY).  
FT METAL 502 502 ZINC (BY SIMILARITY).  
FT METAL 510 510 ZINC (BY SIMILARITY).  
FT DISULFID 399 405 BY SIMILARITY.  
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (BY SIMILARITY).

FT VARSPLIC 428 450 SFVVGSDGYLYQGRGWHWVGAHT -> RLKTKNSFERPLKI  
FT QEVLSLMIL (in isoform 3).  
FT /FTId=VSP\_009079.  
FT Missing (in isoform 3).  
FT /FTId=VSP\_009080.  
FT VARSPLIC 338 366 Missing (in isoform 2).  
FT /FTId=VSP\_009081.  
FT CONFLICT 486 486 MISSING (IN REF. 2).  
SQ SEQUENCE 530 AA; 57706 MW; D3BF52597CE5D1F9 CRC64;  
  
Query Match 5.7%; Score 10; DB 1; Length 530;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RGVHWVGAHT 10  
| | | | | | | | | |  
Db 441 RGVHWVGAHT 450  
  
RESULT 4  
ITN2\_MOUSE  
ID ITN2\_MOUSE STANDARD; PRT; 1658 AA.  
AC Q9Z0R6; Q9Z0R5;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Intersectin 2 (SH3 domain-containing protein 1B) (EH and SH3 domains  
DE protein 2) (EH domain and SH3 domain regulator of endocytosis 2).  
GN ITSN2 OR SH3D1B OR ESE2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=99164083; Pubmed=10064583;  
RA Sengar A.S., Wang W., Bishay J., Cohen S., Egan S.E.;  
RT "The EH and SH3 domain ESE proteins regulate endocytosis by linking to  
RT dynamin and Eps15";  
RL EMBO J. 18:1159-1171 (1999).  
CC -!- FUNCTION: Adapter protein that may provide indirect link between  
CC the endocytic membrane traffic and the actin assembly machinery.  
CC May regulate the formation of clathrin-coated vesicles.  
CC -!- SUBUNIT: Belongs to a complex that may contain multimers of ITSN1,  
CC ITSN2 and Eps15, and different partners according to the step in  
CC the endocytic process.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1; Synonyms=Ese2L, Long;  
CC IsoId=Q9Z0R6-1; Sequence=Displayed;  
CC Name=2; Synonyms=Ese2, Short;  
CC IsoId=Q9Z0R6-2; Sequence=VSP\_003896, VSP\_003897;  
CC -!- TISSUE SPECIFICITY: Widely expressed in adult tissues.  
CC -!- DEVELOPMENTAL STAGE: Widely distributed throughout the adult  
CC forebrain. Prominent expression was observed in the neocortex, the  
CC piriform cortex, the pyramidal cell layers of hippocampus, the  
CC dentate gyrus, in several nuclei of the thalamus and hypothalamus  
CC and in the amygdala.  
CC -!- SIMILARITY: Contains 2 C2 domains.  
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.  
CC -!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.  
CC -!- SIMILARITY: Contains 2 EH domains.  
CC -!- SIMILARITY: Contains 1 PH domain.  
CC -!- SIMILARITY: Contains 5 SH3 domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



CC EMBL; AF132479; AAD19747.1; -.  
DR EMBL; AF132480; AAD19748.1; -.  
DR HSSP; P29355; 1SEM.  
DR MGD; MGI:1338049; Sh3dlb.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CalB.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR000261; EPS15\_homology.  
DR InterPro; IPR000108; Neu\_cyt\_fact\_2.  
DR InterPro; IPR001849; PH.  
DR InterPro; IPR000219; RhogEF.  
DR InterPro; IPR001452; SH3.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00036; ehand; 2.  
DR Pfam; PF00169; PH; 1.  
DR Pfam; PF00621; RhogEF; 1.  
DR Pfam; PF00018; SH3; 5.  
DR PRINTS; PR00499; P67PHOX.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR ProDom; PD000066; SH3; 5.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00054; EFh; 2.  
DR SMART; SM00027; EH; 2.  
DR SMART; SM00233; PH; 1.  
DR SMART; SM00325; RhogEF; 1.  
DR SMART; SM00326; SH3; 5.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS00010; DH\_2; 1.  
DR PROSITE; PS00018; EF\_HAND; 1.  
DR PROSITE; PS00031; EH\_2.  
DR PROSITE; PS00003; PH\_DOMAIN; 1.  
DR PROSITE; PS00002; SH3; 5.  
KW Endocytosis; SH3 domain; Repeat; Coiled coil; Calcium-binding;  
KW Alternative splicing.  
FT DOMAIN 21 109 EH 1.  
FT CA BIND 66 78 EF-HAND (POTENTIAL).  
FT DOMAIN 244 333 EH 2.  
FT DOMAIN 364 716 COILED COIL (POTENTIAL).  
FT DOMAIN 717 778 SH3 1.  
FT DOMAIN 851 909 SH3 2.  
FT DOMAIN 941 999 SH3 3.  
FT DOMAIN 1013 1077 SH3 4.  
FT DOMAIN 1087 1146 SH3 5.  
FT DOMAIN 1169 1356 DH.  
FT DOMAIN 1395 1505 PH.  
FT DOMAIN 1517 1613 C2 DOMAIN.  
FT VARSP LIC 1187 1197 DDLQLVIEVFQ -> GLQLFEQKTL (in isoform 2).  
FT FTID=VSP 003896.  
FT Missing (in isoform 2).  
FT FTID=VSP 003897.  
FT KQ -> NT (IN REF. 1; AAD19748).  
FT CONFLICT 451 452 D -> G (IN REF. 1; AAD19748).  
FT CONFLICT 1161 1161  
SQ SEQUENCE 1658 AA; 188776 MW; 7050EFC5F7983A5 CRC64;

Query Match 4.6%; Score 8; DB 1; Length 1658;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 153 VTSAYAAS 160  
Db 1020 VTSAYAAS 1027

RESULT 5

SECE\_ECOLI STANDARD; PRT; 127 AA.  
ID SECE\_ECOLI  
AC P16920;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Preprotein translocase secE subunit.

GN SECE OR PRLG OR B3981 OR C4936 OR Z5554 OR ECS4904.  
OS Escherichia coli,  
OS Escherichia coli O6, and  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217992, 83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90170882; PubMed=2137819;  
RA Downing W.L., Sullivan S.L., Gottesman M.E., Dennis P.P.;  
RT "Sequence and transcriptional pattern of the essential Escherichia coli secE-nusG operon.";  
RL J. Bacteriol. 172:1621-1627(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89378734; PubMed=2673920;  
RA Schatz P.J., Riggs P.D., Jacq A., Fath M.J., Beckwith J.;  
RT "The secE gene encodes an integral membrane protein required for protein export in Escherichia coli.";  
RL Genes Dev. 3:1035-1044(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MGI655;  
RX MEDLINE=94089392; PubMed=8265357;  
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.;  
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes.";  
RL Nucleic Acids Res. 21:5408-5417(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P., Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D., Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
RN [7]  
RP TOPOLOGY.  
RX MEDLINE=91266903; PubMed=2050112;  
RA Schatz P.J., Bieker K.L., Ottemann K.M., Silhavy T.J., Beckwith J.;  
RT "One of three transmembrane stretches is sufficient for the functioning of the SecE protein, a membrane component of the E. coli secretion machinery.";  
RL EMBO J. 10:1749-1757(1991).  
RN [8]





DR Pfam; PF00584; SecE; 1.  
DR PRINTS; PR01650; SECTRNLCASE.  
DR TIGRFAMs; TIGR00964; 3a0501s06; 1.  
DR PROSITE; PS01067; SECE\_SEC61G; 1.  
KW Protein transpos; Translocation; Transmembrane; Inner membrane;  
KW Complete proteome.  
FT DOMAIN 1 18 CYTOPLASMIC (PROBABLE).  
FT TRANSMEM 19 36 PROBABLE.  
FT DOMAIN 37 44 PERIPLASMIC (PROBABLE).  
FT TRANSMEM 45 63 PROBABLE.  
FT DOMAIN 64 92 CYTOPLASMIC (PROBABLE).  
FT TRANSMEM 93 111 PROBABLE.  
FT DOMAIN 112 127 PERIPLASMIC (PROBABLE).  
SQ SEQUENCE 127 AA; 13689 MW; 640DB5C2080E775D CRC64;  
  
Query Match 4.0%; Score 7; DB 1; Length 127;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 20 VAIVGNY 26  
|||||||  
Db 28 VAIVGNY 34  
  
RESULT 7  
CCMH\_ECOLI STANDARD; PRT; 350 AA.  
AC P33925;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cytochrome c-type biogenesis protein cmh precursor.  
GN CCMH OR B2194 OR SF2278 OR S2408.  
OS Escherichia coli, and  
OS Shigella flexneri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12 / BHB2600;  
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,  
RA Church G.M.;  
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474 (1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
RX MEDLINE=22272406; PubMed=12384590;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.;  
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
RT through comparison with genomes of Escherichia coli K12 and O157.";  
RL Nucleic Acids Res. 30:4432-4441 (2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
RX MEDLINE=22590274; PubMed=12704152;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,

RA Schwartz D.C., Blattner F.R.;  
RT "Complete genome sequence and comparative genomics of Shigella  
RT flexneri serotype 2a strain 2457T.";  
RL Infect. Immun. 71:2775-2786 (2003).  
RN [5]  
RP CHARACTERIZATION, AND GENE NAME.  
RC SPECIES=E.coli;  
RX MEDLINE=95362656; PubMed=7635817;  
RA Thoeny-Meyer L., Fischer F., Kunzler P., Ritz D., Hennecke H.;  
RT "Escherichia coli genes required for cytochrome c maturation.";  
RL J. Bacteriol. 177:4321-4326 (1995).  
CC -!- FUNCTION: May be required for the biogenesis of c-type  
CC cytochromes. Possible subunit of a heme lyase.  
CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).  
CC -!- SIMILARITY: BELONGS TO THE CCMH/CYCL/CCL2/NRFF FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
DR EMBL; U00008; AAA16386.1; -.  
DR EMBL; AE000309; AAC75254.1; -.  
DR EMBL; AE015244; AAN43797.1; -.  
DR EMBL; AE016985; AAP17614.1; -.  
DR PIR; H64988; H64988.  
DR EcoGene; EGI2052; ccmH.  
DR InterPro; IPR005616; CcmH.  
DR InterPro; IPR008940; Prenyl\_trans.  
DR InterPro; IPR001440; TPR.  
DR Pfam; PF03918; CcmH; 1.  
DR ProDom; PD005662; CcmH; 1.  
DR Cytochrome c-type biogenesis; Periplasmic; Heme; Signal;  
KW Complete proteome.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 350 CYTOCHROME C-TYPE BIOGENESIS PROTEIN  
FT  
FT BINDING 43 43 HEME (COVALENT) (POTENTIAL).  
FT BINDING 46 46 HEME (COVALENT) (POTENTIAL).  
SQ SEQUENCE 350 AA; 39089 MW; 002126A1F17F493F CRC64;  
  
Query Match 4.0%; Score 7; DB 1; Length 350;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 63 RQLVRTD 69  
|||||||  
Db 283 RQLVRTD 289  
  
RESULT 8  
MODC\_MYCTU STANDARD; PRT; 369 AA.  
AC P95155; O05126;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Molybdenum import ATP-binding protein modC (EC 3.6.3.29).  
GN MODC OR RV1859 OR MT1907 OR MTCY359.14.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37Rv;  
RA Laqueyrie A.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

```
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McJannet J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwin M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- FUNCTION: Part of the ABC transporter complex modABC (TC 3.A.1.8)
CC involved in molybdenum import. Responsible for energy coupling to
CC the transport system (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + molybdate(Out) = ADP + phosphate
CC + molybdate(In).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Potential).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X99258; CAA67644.1; -.
CC EMBL; Z83859; CAB06128.1; -.
CC EMBL; AE007048; AAK46178.1; -.
CC PIR; C70666; C70666.
CC TIGR; MT1907; -.
CC TubercuList; RV1859; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR008995; MOP_like.
CC InterPro; IPR005116; TOBE.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF03459; TOBE; 1.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC Hydrolase; Transport; Molybdenum; Membrane; ATP-binding;
KW Complete proteome.
KW NP_BIND 31 38 ATP (POTENTIAL).
FT FT 31 111 DAGLVRLGDRVLTDTTEAGVNVVATHDRRVGLLQDPLLPFL
FT FT 51 111 SVAKNVAFGPQCRGMFGSG -> RRLGTFGGPGVDRHRG
FT FT RGECDPRPSSRAAVRPVVVSTPERGQKRGRLRTTTPSRDV
FT FT WVRA (IN REF. 1).
SQ SEQUENCE 369 AA; 38610 MW; D931CC447E70FBD4 CRC64;

Query Match 4.0%; Score 7; DB 1; Length 369;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AGLLRPD 56
Db |||||
45 AGLLRPD 51
```

```
RESULT 9
FLGI_RHIME STANDARD; PRT; 371 AA.
ID FLGI_RHIME
AC Q52948;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Flagellar P-ring protein precursor (Basal body P-ring protein).
GN FLGI OR R00664 OR SMC03032.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RU11/001;
RA Platzner J., Schmitt R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Bounry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: Assembles around the rod to form the L-ring and probably
CC protects the motor/basal body from shearing forces during
CC rotation.
CC -!- SUBUNIT: The basal body constitutes a major portion of the
CC flagellar organelle and consists of four rings (L,P,S, and M)
CC mounted on a central rod (By similarity).
CC -!- SUBCELLULAR LOCATION: Periplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the flgi family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L49337; AAB81415.1; -.
CC EMBL; AL591784; CAC45236.1; -.
CC HAMAP; MF_00416; -.
CC InterPro; IPR001782; Flag_Flgi.
CC Pfam; PF02119; Flgi; 1.
CC PRINTS; PR01010; FLGPRINGFLGI.
KW Flagellum; Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 371 FLAGELLAR P-RING PROTEIN.
FT CONFLICT 7 7 N -> K (IN REF. 1).
FT CONFLICT 15 15 V -> A (IN REF. 1).
FT CONFLICT 323 323 G -> GR (IN REF. 1).
SQ SEQUENCE 371 AA; 38357 MW; 06BB4E95EE029100 CRC64;

Query Match 4.0%; Score 7; DB 1; Length 371;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TSAYAAS 160
Db |||||
21 TSAYAAS 27

RESULT 10
```



```
YEHY_ECOLI
ID_YEHY_ECOLI STANDARD; PRT; 385 AA.
AC P33361; P76435;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical ABC transporter permease protein yehY.
GN YEHY OR B2130.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / BHB2600;
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA Church G.M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
CC SYSTEM. PROBABLY RESPONSIBLE FOR THE TRANSLOCATION OF THE
CC SUBSTRATE ACROSS THE MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family. CysTW subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00007; AAA60493.1; ALT_INIT.
CC EMBL; AE000302; AAC75191.1; -.
CC PIR; A64981; A64981.
CC EcoGene; EG12011; yehY.
CC InterPro; IPR000515; BPD_transp.
CC Pfam; PF00528; BPD_transp; 1.
CC PROSITE; PS50928; ABC_TM1; 1.
CC KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
CC Complete proteome.
CC FT TRANSMEM 6 26 POTENTIAL.
CC FT TRANSMEM 45 65 POTENTIAL.
CC FT TRANSMEM 70 90 POTENTIAL.
CC FT TRANSMEM 107 127 POTENTIAL.
CC FT TRANSMEM 181 201 POTENTIAL.
CC FT TRANSMEM 225 245 POTENTIAL.
CC FT TRANSMEM 248 268 POTENTIAL.
CC FT TRANSMEM 323 343 POTENTIAL.
CC FT TRANSMEM 347 367 POTENTIAL.
CC FT CONFLICT 346 385 LLSSAIDLVLGVIPVIVLAVLTDALEFDLLIALLKVKRND
CC -> AAKQRH (IN REF. 1).
CC SQ SEQUENCE 385 AA; 41138 MW; E57055E3A2B141CC CRC64;

Query Match 4.0%; Score 7; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 DALEFDLL 79
| | | | |
Db 369 DALEFDLL 375
```

```
RESULT 11
DEOB_BACHD
ID_DEOB_BACHD STANDARD; PRT; 393 AA.
AC Q9KCN9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphotomutase (EC 5.4.2.7) (Phosphodeoxyribomutase).
GN DEOB OR DRM OR BH1530.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- FUNCTION: Phosphotransfer between the C1 and C5 carbon atoms of
CC pentose (By similarity).
CC -!- CATALYTIC ACTIVITY: D-ribose 1-phosphate = D-ribose 5-phosphate.
CC -!- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 1-phosphate = 2-deoxy-D-
CC ribose 5-phosphate.
CC -!- COFACTOR: Binds 1 or 2 manganese ions (Potential).
CC -!- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the phosphotomutase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001512; BAB05249.1; -.
CC PIR; B83841; B83841.
CC HAMAP; MF_00740; -.
CC InterPro; IPR006124; Metalloenzyme.
CC Pfam; PF01676; Metalloenzyme; 1.
CC KW Isomerase; Metal-binding; Manganese; Complete proteome.
CC FT METAL 15 15 MANGANESE (BY SIMILARITY).
CC FT METAL 293 293 MANGANESE (BY SIMILARITY).
CC FT METAL 329 329 MANGANESE (BY SIMILARITY).
CC FT METAL 330 330 MANGANESE (BY SIMILARITY).
CC FT METAL 341 341 MANGANESE (BY SIMILARITY).
CC SQ SEQUENCE 393 AA; 43589 MW; 9CA37EDACF8E544A CRC64;

Query Match 4.0%; Score 7; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GAHTLGH 13
| | | | |
Db 32 GAHTLGH 38

RESULT 12
DEOB_BACSU
ID_DEOB_BACSU STANDARD; PRT; 394 AA.
AC P46353;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphotomutase (EC 5.4.2.7) (Phosphodeoxyribomutase).
GN DRM OR BSU23500.
```

OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BR151;  
RX MEDLINE=20005610; PubMed=10537218;  
RA Schuch R., Garibian A., Saxild H.H., Piggot P.J., Nygaard P.;  
RT "Nucleosides as a carbon source in Bacillus subtilis: characterization  
RT of the drm-pupG operon.";  
RL Microbiology 145:2957-2966(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / JH642;  
RX MEDLINE=97124195; PubMed=8969508;  
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
RA Kobayashi Y.;  
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
RT the Bacillus subtilis genome containing the skin element and many  
RT sporulation genes.";  
RL Microbiology 142:3103-3111(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell I.B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T., Yata K.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis.";  
RL Nature 390:249-256(1997).  
CC -!- FUNCTION: Phosphotransfer between the C1 and C5 carbon atoms of  
CC pentose.  
CC -!- CATALYTIC ACTIVITY: D-ribose 1-phosphate = D-ribose 5-phosphate.  
CC -!- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 1-phosphate = 2-deoxy-D-  
CC ribose 5-phosphate.  
CC -!- COFACTOR: Binds 1 or 2 manganese ions (Potential).  
CC -!- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the phosphopentomutase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; U32685; AAA74433.1; -.  
DR EMBL; D84432; BAA12650.1; -.  
DR EMBL; Z99116; CAB14282.1; -.  
DR PIR; B69619; B69619.  
DR Subtilist; BG11331; drm.  
DR HAMAP; MF\_00740; -. 1.  
DR InterPro; IPR006124; Metalloenzyme.  
DR Pfam; PF01676; Metalloenzyme; 1.  
KW Isomerase; Metal-binding; Manganese; Complete proteome.  
FT METAL 15 15 MANGANESE (BY SIMILARITY).  
FT METAL 293 293 MANGANESE (BY SIMILARITY).  
FT METAL 329 329 MANGANESE (BY SIMILARITY).  
FT METAL 330 330 MANGANESE (BY SIMILARITY).  
FT METAL 341 341 MANGANESE (BY SIMILARITY).  
FT CONFLICT 78 78 L -> M (IN REF. 1).  
FT CONFLICT 207 207 Q -> QFK (IN REF. 1).  
FT CONFLICT 316 316 G -> E (IN REF. 1).  
FT SEQUENCE 394 AA; 43929 MW; 390E7C9A1E4B524D CRC64;  
SQ  
Query Match 4.0%; Score 7; DB 1; Length 394;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 GAHTLGH 13  
Db 32 GAHTLGH 38  
RESULT 13  
PILC\_PSEPU STANDARD; PRT; 401 AA.  
ID PILC\_PSEPU STANDARD; PRT; 401 AA.  
AC P36641;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Type 4 fimbrial assembly protein pilC.  
GN pilC.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WCS358;  
RX MEDLINE=94131942; PubMed=7905475;  
RA de Groot A., Heijnen I., de Cock H., Filloux A., Tommassen J.;  
RT "Characterization of type IV pilus genes in plant growth-promoting  
RT Pseudomonas putida WCS358.";  
RL J. Bacteriol. 176:642-650(1994).  
CC -!- FUNCTION: INVOLVED IN THE TRANSLLOCATION OF THE TYPE IV PILIN  
CC (PILA).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Probable).  
CC -!- SIMILARITY: BELONGS TO THE PULF/OUTF/EXEF/XPSF/XCPS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X74276; CAA52333.1; -.  
DR PIR; A36961; A36961.  
DR InterPro; IPR003004; Bac\_GSPF.  
DR InterPro; IPR001992; Bact\_secr\_systII.  
DR Pfam; PF00482; GSPII\_F; 1.  
DR PRINTS; PR00812; BCTERIALGSPF.  
DR PROSITE; PS00874; T2SP\_F; 1.  
KW Transport; Transmembrane; Inner membrane; Fimbria.  
FT TRANSMEM 166 186 POTENTIAL.



```
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 375 395 POTENTIAL.
SQ SEQUENCE 401 AA; 43084 MW; DD5B4F6123DFCE65 CRC64;

Query Match 4.0%; Score 7; DB 1; Length 401;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 VRAGLLR 54
   |||||
Db 32 VRAGLLR 38

RESULT 14
UL27 HCMVA STANDARD; PRT; 608 AA.
ID UL27 HCMVA
AC P16763;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical protein UL27.
GN UL27.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kourazides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-6 AND
CC HSV-7 U4, HSV-6 AND HSV-7 U5 AND HCMV UL27.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; X17403; CAA35426.1; -.
CC PIR; S09790; S09790.
CC Hypothetical protein.
KW SEQUENCE 608 AA; 69220 MW; 81225A5D00E2980E CRC64;

Query Match 4.0%; Score 7; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AVRAGLL 53
   |||||
Db 69 AVRAGLL 75

RESULT 15
DD15 SCHPO STANDARD; PRT; 735 AA.
ID DD15 SCHPO
AC O42945;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative pre-mRNA splicing factor ATP-dependent RNA helicase
DE C16H5.10C.
GN SPBC16H5.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
```

```
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Pre-mRNA processing factor involved in disassembly of
CC spliceosomes after the release of mature mRNA (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DEAH
CC SUBFAMILY. DDX15/PRP43 ORTHOLOG.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AL022104; CAA17908.1; -.
CC PIR; T39615; T39615.
CC GenedB_SPombe; SPBC16H5.10C; -.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR002464; DEAH_box.
CC InterPro; IPR001650; Helicase_C.
CC InterPro; IPR007502; Helicase_dom.
CC Pfam; PF04408; HA2; 1.
CC Pfam; PF00271; helicase_C; 1.
CC SMART; SM00487; DEXDc; 1.
CC SMART; SM00490; HELICc; 1.
CC PROSITE; PS00690; DEAH_ATP_HELICASE; 1.
KW Hypothetical protein; mRNA processing; mRNA splicing; Helicase;
KW ATP-binding; Nuclear protein.
FT NP_BIND 98 105 ATP (POTENTIAL).
FT SITE 197 200 DEAH BOX.
SQ SEQUENCE 735 AA; 83803 MW; DC2A25F145F5A5C4 CRC64;

Query Match 4.0%; Score 7; DB 1; Length 735;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 EAALRTV 39
   |||||
Db 270 EAALRTV 276
```

```
RESULT 16
AOC3_MOUSE
ID_AOC3_MOUSE STANDARD; PRT; 765 AA.
AC O70423;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Membrane copper amine oxidase (EC 1.4.3.6) (Vascular adhesion protein-1) (VAP-1).
DE AOC3 OR VAP1.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c, and 129/SvJ;
RX MEDLINE=98414290; PubMed=9743358;
RA Bono P., Salmi M., Smith D.J., Leppanen I., Horelli-Kuitunen N.,
RA Palotie A., Jalkanen S.;
RT "Isolation, structural characterization, and chromosomal mapping of the mouse vascular adhesion protein-1 gene and promoter.";
RL J. Immunol. 161:2953-2960(1998).
CC -!- FUNCTION: Cell adhesion protein that participate in lymphocyte recirculation by mediating the binding of lymphocytes to peripheral lymph node vascular endothelial cells in an I-selectin-independent fashion. Has a monoamine oxidase activity.
CC -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 topaquinone per subunit.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- PTM: Topaquinone (TPQ) is generated by copper-dependent autoxidation of a specific tyrosyl residue (By similarity).
CC -!- PTM: N- and O-glycosylated (By similarity).
CC -!- SIMILARITY: Belongs to the copper/topaquinone oxidase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR EMBL; AF054831; AAC23747.1; -.
DR EMBL; AF078705; AAC35839.1; -.
DR MGD; MGI:1306797; Aoc3.
DR InterPro; IPR000269; CuNH_oxidase.
DR Pfam; PF01179; Cu_amine_oxid; 1.
DR Pfam; PF02727; Cu_amine_oxidn2; 1.
DR Pfam; PF02728; Cu_amine_oxidn3; 1.
DR PRINTS; PR00766; CUDAOXIDASE.
DR PROSITE; PS01164; COPPER_AMINE_OXID_1; 1.
DR PROSITE; PS01165; COPPER_AMINE_OXID_2; 1.
DR Oxidoreductase; Copper; TPQ; Glycoprotein; Transmembrane;
KW Signal-anchor; Cell adhesion; Metal-binding.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 27 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
FT FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 28 765
FT MOD RES 471 471 TOPAQUINONE (BY SIMILARITY).
FT METAL 520 520 COPPER (POTENTIAL).
FT METAL 522 522 COPPER (POTENTIAL).
FT METAL 684 684 COPPER (POTENTIAL).
FT BINDING 673 683 AMILORIDE (BY SIMILARITY).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 592 592 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 765 AA; 84533 MW; 7489ED67D3DBB44D CRC64;
```

```
Query Match 4.0%; Score 7; DB 1; Length 765;
Best Local Similarity 100.0%; Pred.No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 VGAHTLG 12
Db 511 VGAHTLG 517

RESULT 17
ECA2_ARATH STANDARD; PRT; 1054 AA.
ID_ECA2_ARATH
AC O23087;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Calcium-transporting ATPase 2, endoplasmic reticulum-type (EC 3.6.3.8).
GN ECA2 OR ACA5 OR AT4G00900 OR A_TM018A10.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99365310; PubMed=10433975;
RA Pittman J.K., Mills R.F., O'Connor C.D., Williams L.E.;
RT "Two additional type IIA Ca(2+)-ATPases are expressed in Arabidopsis thaliana: evidence that type IIA sub-groups exist.";
RL Gene 236:137-147(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N., Harris B., Ansoerge W., Brandt P., Grivell L.A., Rieger M., Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M., Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T., Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F., Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E., Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M., Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W., Mooijman P., Klein lankhorst R., Rose M., Hauf J., Koeter P., Bernseiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlay K., Mayes R., Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A., Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E., Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R., Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S., Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E., Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T., Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C., Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S., Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K., Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L., Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J., Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
```



RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,  
RA Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,  
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
thaliana.";  
RL Nature 402:769-777(1999).  
CC -!- FUNCTION: This magnesium-dependent enzyme catalyzes the hydrolysis  
CC of ATP coupled with the translocation of calcium from the cytosol  
CC to an endomembrane compartment.  
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+)(Cis) = ADP + phosphate +  
CC Ca(2+)(Trans).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type  
CC ATPases). Subfamily IIA.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
-----  
DR EMBL; AJ132387; CAA10659.1; -.  
DR EMBL; AF013294; AAB62850.1; -.  
DR EMBL; AL161472; CAB80899.1; -.  
DR PIR; T01556; T01556.  
DR HSSP; P04191; 1EUL.  
DR InterPro; IPR001757; ATPase\_E1-E2.  
DR InterPro; IPR005782; Calcium\_ATPase.  
DR InterPro; IPR006069; Cation\_ATPase.  
DR InterPro; IPR006068; Cation\_ATPase\_C.  
DR InterPro; IPR004014; Cation\_ATPase\_N.  
DR InterPro; IPR008250; E1-E2\_ATPase\_reg.  
DR InterPro; IPR005834; Hydrolase.  
DR Pfam; PF00689; Cation\_ATPase\_C; 1.  
DR Pfam; PF00690; Cation\_ATPase\_N; 1.  
DR Pfam; PF00122; E1-E2\_ATPase; 1.  
DR Pfam; PF00702; Hydrolase; 1.  
DR PRINTS; PR00119; CATATPASE.  
DR PRINTS; PR00121; NAKATPASE.  
DR TIGRFAMs; TIGR01116; ATPase-IIA1\_Ca; 1.  
DR TIGRFAMs; TIGR01494; ATPase\_P-type; 8.  
DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;  
KW ATP-binding; Metal-binding; Magnesium; Calcium-binding;  
KW Multigene family.  
FT DOMAIN 1 53 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 54 74 POTENTIAL.  
FT DOMAIN 75 98 LUMENAL (POTENTIAL).  
FT TRANSMEM 99 118 POTENTIAL.  
FT DOMAIN 119 262 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 263 282 POTENTIAL.  
FT DOMAIN 283 312 LUMENAL (POTENTIAL).  
FT TRANSMEM 313 330 POTENTIAL.  
FT DOMAIN 331 782 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 783 802 POTENTIAL.  
FT DOMAIN 803 812 LUMENAL (POTENTIAL).  
FT TRANSMEM 813 833 POTENTIAL.  
FT DOMAIN 834 853 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 854 876 POTENTIAL.  
FT DOMAIN 877 949 LUMENAL (POTENTIAL).  
FT TRANSMEM 950 969 POTENTIAL.  
FT DOMAIN 970 982 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 983 1001 POTENTIAL.  
FT DOMAIN 1002 1016 LUMENAL (POTENTIAL).  
FT TRANSMEM 1017 1037 POTENTIAL.  
FT DOMAIN 1038 1054 CYTOPLASMIC (POTENTIAL).  
FT MOD\_RES 368 PHOSPHORYLATION (BY SIMILARITY).

FT METAL 727 727 MAGNESIUM (BY SIMILARITY).  
FT METAL 731 731 MAGNESIUM (BY SIMILARITY).  
FT METAL 321 321 CALCIUM 2 (VIA CARBONYL OXYGEN) (BY  
FT METAL 322 322 SIMILARITY).  
FT METAL 322 322 CALCIUM 2 (VIA CARBONYL OXYGEN) (BY  
FT METAL 324 324 SIMILARITY).  
FT METAL 324 324 CALCIUM 2 (VIA CARBONYL OXYGEN) (BY  
FT METAL 326 326 SIMILARITY).  
FT METAL 793 793 CALCIUM 2 (BY SIMILARITY).  
FT METAL 796 796 CALCIUM 1 (BY SIMILARITY).  
FT METAL 821 821 CALCIUM 1 (BY SIMILARITY).  
FT METAL 824 824 CALCIUM 2 (BY SIMILARITY).  
FT METAL 825 825 CALCIUM 1 (BY SIMILARITY).  
FT METAL 825 825 CALCIUM 1 (BY SIMILARITY).  
FT METAL 960 960 CALCIUM 2 (BY SIMILARITY).  
SQ SEQUENCE 1054 AA; 115829 MW; B0D7F7237A3496AF CRC64;  
  
Query Match 4.0%; Score 7; DB 1; Length 1054;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 30 LPTEAAL 36  
| | | | |  
Db 449 LPTEAAL 455  
  
RESULT 18  
SEN1 SCHPO STANDARD; PRT; 1687 AA.  
ID SEN1 SCHPO STANDARD; PRT; 1687 AA.  
AC Q92355;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Endonuclease sen1.  
GN SEN1 OR SPAC6G9.10C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]\_SEQUENCE FROM N.A.  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu B., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
RN [2]  
RP SEQUENCE OF 1230-1241 AND 1489-1493, FUNCTION, AND SUBUNIT.

```
RC STRAIN=972;
RX MEDLINE=20014567; PubMed=10545196;
RA Kim H.-D., Choe J., Seo Y.-S.;
RT "The senl(+) gene of Schizosaccharomyces pombe, a homologue of budding
RL Biochemistry 38:14697-14710(1999).
CC -!- FUNCTION: Binds to DNA and RNA and has a 5'->3' endonuclease
CC activity. Has a role in tRNA, mRNA, and snRNA splicing.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z81317; CAB03612.1; -.
DR PIR; T39072; T39072.
DR GeneDB SPombe; SPAC6G9.10c; -.
KW tRNA processing; mRNA splicing; Nuclear protein; ATP-binding;
KW Hydrolase; Nuclease.
FT NP BIND 1152 1159 ATP (BY SIMILARITY).
SQ SEQUENCE 1687 AA; 192547 MW; 92C82F2049E3680D CRC64;

Query Match 4.0%; Score 7; DB 1; Length 1687;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 DAIFDLL 79
Db 356 DAIFDLL 362

RESULT 19
NDL DROME STANDARD; PRT; 2616 AA.
AC P98I59;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serine protease nudel precursor (EC 3.4.21.-).
GN NDL.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Ovary;
RX MEDLINE=95401268; PubMed=7671306;
RA Hong C.C., Hashimoto C.;
RT "An unusual mosaic protein with a protease domain, encoded by the
RT nudel gene, is involved in defining embryonic dorsoventral polarity
RT in Drosophila.";
RL Cell 82:785-794(1995).
CC -!- FUNCTION: Nudel, Pipe and Windbeutel together trigger the protease
CC cascade within the extraembryonic perivitelline compartment which
CC induces dorsoventral polarity of the Drosophila embryo. Nudel is
CC directly involved in locally producing the Toll ligand.
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC -!- TISSUE SPECIFICITY: Follicle.
CC -!- PTM: Requires cleavage for activation (presumably).
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 11 LDL-receptor class A domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U29153; AAA83086.1; -.
DR PIR; A57096; A57096.
DR HSP; P00763; LDPO.
DR MEROPS; S01.013; -.
DR FlyBase; FBgn0002926; ndl.
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR GO; GO:0007306; P:insect chorion formation; IMP.
DR GO; GO:0007310; P:oocyte dorsal/ventral axis determination; NAS.
DR GO; GO:0016485; P:protein processing; IGI.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00057; ldl_recept_a; 6.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00192; LDLa; 8.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; 6.
DR PROSITE; PS50068; LDLRA_2; 8.
DR PROSITE; PS50240; TRYPSIN_DOM; 2.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Serine protease; Glycoprotein; Signal; Developmental protein;
KW Hydrolase; Repeat; Zymogen; Extracellular matrix.
FT SIGNAL 1 43 POTENTIAL.
FT CHAIN 44 2616 SERINE PROTEASE NUDEL.
FT REPEAT 261 269 WIID 1.
FT REPEAT 320 328 WIID 2.
FT REPEAT 399 407 WIID 3.
FT REPEAT 446 454 WIID 4.
FT REPEAT 477 485 WIID 5.
FT REPEAT 528 536 WIID 6.
FT DOMAIN 889 929 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 929 956 LDL-RECEPTOR CLASS A 2 (PARTIAL).
FT DOMAIN 955 1006 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 1145 1383 SERINE PROTEASE 1.
FT DOMAIN 1394 1432 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 1489 1702 SER/THR-RICH.
FT DOMAIN 1713 1743 LDL-RECEPTOR CLASS A 5 (PARTIAL).
FT DOMAIN 1745 1775 LDL-RECEPTOR CLASS A 6 (PARTIAL).
FT DOMAIN 1774 1813 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 2027 2301 SERINE PROTEASE 2.
FT DOMAIN 2308 2346 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 2349 2389 LDL-RECEPTOR CLASS A 9.
FT DOMAIN 2387 2419 LDL-RECEPTOR CLASS A 10 (PARTIAL).
FT DOMAIN 2419 2459 LDL-RECEPTOR CLASS A 11.
FT SITE 1031 1033 CELL ATTACHMENT SITE (POTENTIAL).
FT ACT_SITE 1185 1185 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1233 1233 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1332 1332 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 891 905 BY SIMILARITY.
FT DISULFID 899 918 BY SIMILARITY.
FT DISULFID 912 927 BY SIMILARITY.
FT DISULFID 957 982 BY SIMILARITY.
FT DISULFID 964 995 BY SIMILARITY.
FT DISULFID 989 1004 BY SIMILARITY.
FT DISULFID 1170 1186 BY SIMILARITY.
FT DISULFID 1276 1338 POTENTIAL.
FT DISULFID 1305 1317 BY SIMILARITY.
FT DISULFID 1328 1359 BY SIMILARITY.
FT DISULFID 1396 1408 BY SIMILARITY.
FT DISULFID 1401 1421 BY SIMILARITY.
FT DISULFID 1415 1430 BY SIMILARITY.
FT DISULFID 1776 1789 BY SIMILARITY.
FT DISULFID 1783 1802 BY SIMILARITY.
FT DISULFID 1796 1811 BY SIMILARITY.
```



FT DISULFID 2310 2320 BY SIMILARITY.  
FT DISULFID 2315 2333 BY SIMILARITY.  
FT DISULFID 2327 2344 BY SIMILARITY.  
FT DISULFID 2351 2364 BY SIMILARITY.  
FT DISULFID 2358 2377 BY SIMILARITY.  
FT DISULFID 2371 2387 BY SIMILARITY.  
FT DISULFID 2421 2435 BY SIMILARITY.  
FT DISULFID 2428 2448 BY SIMILARITY.  
FT DISULFID 2442 2457 BY SIMILARITY.  
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 417 417 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 492 492 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 794 794 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).  
FT CARBOHYD 827 827 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 829 829 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).  
FT CARBOHYD 861 861 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 975 975 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1064 1064 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1878 1878 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1956 1956 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2023 2023 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2144 2144 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2173 2173 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2197 2197 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2237 2237 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2269 2269 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2420 2420 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2556 2556 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2601 2601 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 2616 AA; 292371 MW; 25DCB13213DC7D13 CRC64;

Query Match 4.0%; Score 7; DB 1; Length 2616;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 VRAGLLR 54  
Db 1197 VRAGLLR 1203

RESULT 20  
Y872\_VIBCH  
ID Y872\_VIBCH STANDARD; PRT; 67 AA.  
AC Q9KTN0;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical UPF0253 protein VC0872.  
GN VC0872.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=566;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=El Tor N16961 / Serotype O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae.";  
RL Nature 406:477-483(2000).  
CC -!- SIMILARITY: Belongs to the UPF0253 family.

CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC

DR EMBL; AE004172; AAF94034.1; -.  
DR PIR; E82268; E82268.  
DR TIGR; VC0872; -.  
DR HAMAP; MF\_01064; -; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 67 AA; 7179 MW; 0FB8590063BFE42E CRC64;

Query Match 3.4%; Score 6; DB 1; Length 67;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 SCAVRA 50  
Db 30 SCAVRA 35

RESULT 21  
NXT1\_DROME  
ID NXT1\_DROME STANDARD; PRT; 133 AA.  
AC Q9V3H8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE NTF2-related export protein (p15).  
GN NXT1 OR CG12752.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20036817; PubMed=10567585;  
RA Black B.E., Levesque L., Holaska J.M., Wood T.C., Paschal B.M.;  
RA "Identification of an NTF2-related factor that binds Ran-GTP and  
RA regulates nuclear protein export.";  
RL Mol. Cell. Biol. 19:8616-8624(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Flosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
CC -!- FUNCTION: Stimulator of protein export for NES-containing  
CC proteins. Also plays a role in the nuclear export of U1 snRNA,  
CC tRNA, and mRNA (By similarity).  
CC -!- SUBUNIT: Preferentially binds Ran-GTP (By similarity). Associates  
CC with NXF1.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -!- SIMILARITY: Contains 1 NTF2 domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF156959; AAD54944.1; -.  
DR EMBL; AE003462; AAF47066.1; -.  
DR FlyBase; FBgn0028411; Nxt1.  
DR GO; GO:0016973; P:poly(A)+ mRNA-nucleus export; IMP.  
DR InterPro; IPR002075; NTF2.  
DR Pfam; PF02136; NTF2; 1.  
DR PROSITE; PS50177; NTF2\_DOMAIN; 1.  
DR Transport; Protein transport; mRNA transport; Nuclear protein.  
KW DOMAIN 15 130 NTF2.  
FT SEQUENCE 133 AA; 15181 MW; C6E664950AA370AA CRC64;  
SQ  
Query Match 3.4%; Score 6; DB 1; Length 133;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 116 SCARTA 121  
Db 11 SCARTA 16  
RESULT 22  
NRDI\_MYCLE  
ID NRDI\_MYCLE STANDARD; PRT; 138 AA.  
AC Q9CBP9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Nrdi protein.  
GN NRDI OR MLI735.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TN;  
RX MEDLINE=21128732; PubMed=11234002;  
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,

RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
RA Barrell B.G.;  
RT "Massive gene decay in the leprosy bacillus."  
RL Nature 409:1007-1011(2001).  
CC -!- FUNCTION: Not known; probably involved in ribonucleotide reductase  
CC function.  
CC -!- SIMILARITY: Belongs to the nrdI family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AL583923; CAC30688.1; -.  
DR PIR; A87126; A87126.  
DR Leproma; ML1735; -.  
DR HAMAP; MF\_00128; -; 1.  
DR InterPro; IPR004465; NrdI.  
DR TIGRFAMs; TIGR00333; nrdI; 1.  
KW Complete proteome.  
SQ SEQUENCE 138 AA; 15166 MW; 9E911349F484E916 CRC64;  
Query Match 3.4%; Score 6; DB 1; Length 138;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 47 AVRAGL 52  
Db 126 AVRAGL 131  
RESULT 23  
FER1\_MESCR  
ID FER1\_MESCR STANDARD; PRT; 148 AA.  
AC O04683;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ferredoxin I, chloroplast precursor.  
OS Mesembryanthemum crystallinum (Common ice plant).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Aizoaceae; Mesembryanthemum.  
OX NCBI\_TaxID=3544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Michalowski C.B., Bohnert H.J.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Ferredoxins are iron-sulfur proteins that transfer  
CC electrons in a wide variety of metabolic reactions.  
CC -!- COFACTOR: Binds 1 2Fe-2S cluster.  
CC -!- SUBCELLULAR LOCATION: Chloroplast.  
CC -!- SIMILARITY: Belongs to the 2Fe2S plant-type ferredoxin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF003125; AAB61593.1; -.  
DR PIR; T12417; T12417.  
DR HSSP; P00221; 1A70.  
DR InterPro; IPR006057; 2Fe2S.  
DR InterPro; IPR006058; 2Fe2S\_fd\_BS.  
DR InterPro; IPR001041; Ferredoxin.



```
DR Pfam; PF00111; fer2; 1.
DR PRINTS; PR00159; 2FE2SFERDOXIN.
DR PROSITE; PS00197; 2FE2S FERREDOXIN; 1.
KW Electron transport; Metal-binding; Iron-sulfur; Iron; 2Fe-2S;
KW Chloroplast; Transit peptide.
FT TRANSIT 1 51 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 52 148 FERREDOXIN I.
FT METAL 90 90 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 95 95 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 128 128 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 148 AA; 15358 MW; B3C974ECFF0075A0 CRC64;

Query Match
Best Local Similarity 3.4%; Score 6; DB 1; Length 148;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 TAALPT 32
Db 25 TAALPT 30

RESULT 24
RL15 THETH STANDARD; PRT; 150 AA.
AC P74910;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L15.
GN RPLO OR RPL15.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VK1;
RX MEDLINE=97390124; PubMed=9249063;
RA Vysotskaya V.S., Scherbakov D.V., Garber M.B.;
RT "Sequencing and analysis of the Thermus thermophilus ribosomal protein
  gene cluster equivalent to the spectinomycin operon.";
RL Gene 193:23-30(1997).
CC -!- FUNCTION: This protein binds directly to 23S ribosomal RNA (By
  similarity).
CC -!- SIMILARITY: Belongs to the L15P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X90765; CAA62292.1; -.
DR InterPro; IPR001196; Ribosomal_L15.
DR InterPro; IPR005749; Ribosomal_L15_b.
DR Pfam; PF00256; L15; 1.
DR Pfam; PF01305; Ribosomal_L15; 1.
DR TIGRFAMs; TIGR01071; rplO_bact; 1.
DR PROSITE; PS00475; RIBOSOMAL_L15; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 150 AA; 16281 MW; 0175AF33F530AD4F CRC64;

Query Match
Best Local Similarity 3.4%; Score 6; DB 1; Length 150;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 VRAGLL 53
Db 101 VRAGLL 106
```

```
RESULT 25
DRPD CRAPL STANDARD; PRT; 151 AA.
AC P22241;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Desiccation-related protein PCC27-45.
OS Craterostigma plantagineum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiales; Lamiales incertae sedis; Lindernieae;
OC Craterostigma.
OX NCBI_TaxID=4153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Piatkowski D., Schneider K., Salamini F., Bartels D.;
RT "Characterization of five abscisic acid-responsive cDNA clones
  isolated from the desiccation-tolerant plant Craterostigma
  plantagineum and their relationship to other water-stress genes.";
RL Plant Physiol. 94:1682-1688(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94035172; PubMed=8220473;
RA Michel D., Salamini F., Bartels D., Dale P., Baga M., Szalay A.;
RT "Analysis of a desiccation and ABA-responsive promoter isolated from
  the resurrection plant Craterostigma plantagineum.";
RL Plant J. 4:29-40(1993).
CC -!- INDUCTION: By desiccation (leaves) and by abscisic acid (ABA)
  (leaves and callus).
CC -!- SIMILARITY: BELONGS TO THE LEA TYPE 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62990; AAA63615.1; -.
DR EMBL; X69883; CAA49510.1; -.
DR PIR; S29970; S29970.
DR InterPro; IPR004864; LEA_2.
DR Pfam; PF03168; LEA_2; 1.
DR ProDom; PD010978; LEA_2; 1.
SQ SEQUENCE 151 AA; 16269 MW; 28A0521541905689 CRC64;

Query Match
Best Local Similarity 3.4%; Score 6; DB 1; Length 151;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 GNIPDP 152
Db 74 GNIPDP 79

RESULT 26
Y087 DEIRA
ID Y087 DEIRA STANDARD; PRT; 157 AA.
AC Q9RY63;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0168 protein DR0087.
GN DR0087.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -!- SIMILARITY: Belongs to the UPF0168 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001872; AAF09680.1; -.
DR PIR; G75560; G75560.
DR TIGR; DR0087; -.
DR HAMAP; MF 00440; -.
DR InterPro; IPR005144; ATP.
DR InterPro; IPR003796; DUF193.
DR Pfam; PF03477; ATP-cone; 1.
DR TIGRFAMs; TIGR00244; TIGR00244; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 157 AA; 17996 MW; 5207A871556F8673 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 EAALRT 38
Db 81 EAALRT 86

RESULT 27
APT_STRCL STANDARD; PRT; 178 AA.
AC Q93AJ8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN APT.
OS Streptomyces clavuligerus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1901;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27064 / DSM 738 / NRRL 3585;
RA Jin W., Kim J.Y., Lee K.J.;
RT "Cloning and characterization of a ppGpp synthetase gene (relA) of
RT Streptomyces clavuligerus ATCC27064.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes a salvage reaction resulting in the formation
CC of AMP, that is energetically less costly than de novo synthesis.
CC -!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: Purine salvage.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the purine/pyrimidine
CC phosphoribosyltransferase family.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF421216; AAL16894.1; -.
DR HAMAP; MF 00004; -.
DR InterPro; IPR005764; Ade_phospho_trans.
DR InterPro; IPR002375; Pr/py_rp_transf.
DR InterPro; IPR000836; PRTransferase.
DR Pfam; PF00156; Priboyltran; 1.
DR TIGRFAMs; TIGR01090; apt; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage.
SQ SEQUENCE 178 AA; 18469 MW; E8D36FD6ABB25C69 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AVRAGL 52
Db 73 AVRAGL 78

RESULT 28
RS5_THEME STANDARD; PRT; 178 AA.
AC Q9XIJ2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 30S ribosomal protein S5.
GN RPSE OR TM1483.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- FUNCTION: With S4 and S12 plays an important role in translational
CC accuracy (By similarity).
CC -!- FUNCTION: Located at the back of the 30S subunit body where it
CC stabilizes the conformation of the head with respect to the body
CC (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S4
CC and S8 (By similarity).
CC -!- DOMAIN: The N-terminal domain interacts with the head of the 30S
CC subunit; the C-terminal domain interacts with the body and
CC contacts protein S4. The interaction surface between S4 and S5 is
CC involved in control of translational fidelity.
CC -!- SIMILARITY: Contains 1 S5 DRBM domain.
CC -!- SIMILARITY: Belongs to the S5P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```



```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001798; AAD36549.1; -.
DR PIR; C72248; C72248.
DR HSSP; P02357; LPKP.
DR TIGR; TM1483; -.
DR HAMAP; MF_01307; -; 1.
DR InterPro; IPR000851; Ribosomal_S5.
DR InterPro; IPR005712; Ribosomal_S5_b/o.
DR InterPro; IPR005324; Ribosomal_S5_C.
DR Pfam; PF00333; Ribosomal_S5; 1.
DR Pfam; PF03719; Ribosomal_S5_C; 1.
DR TIGRFAMS; TIGR01021; rpsE_bact; 1.
DR PROSITE; PS00585; RIBOSOMAL_S5; 1.
DR PROSITE; PS50881; S5_DSRHD; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
FT DOMAIN 15 78 S5_DRBM.
SQ SEQUENCE 178 AA; 19125 MW; E7153790C3DAE18F CRC64;

Query Match 3.4%; Score 6; DB 1; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGN 25
Db 40 VAIVGN 45

RESULT 29
APT_STRGB STANDARD; PRT; 182 AA.
AC Q8KIQ0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN APT.
OS Streptomyces galbus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33898;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 40480;
RA Wehmeier U.F.;
RT "Sequence of the genes secD, secF and apt from Streptomyces galbus DSM40480.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes a salvage reaction resulting in the formation of AMP, that is energetically less costly than de novo synthesis.
CC -!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: Purine salvage.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the purine/pyrimidine phosphoribosyltransferase family.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
EMBL; AJ505987; CAD44527.1; -.
DR HAMAP; MF_00004; -; 1.
DR InterPro; IPR005764; Ade_phospho_trans.
DR InterPro; IPR002375; Pr/py_rp_transf.
DR Pfam; PF00156; Pribosyltran; 1.
DR TIGRFAMS; TIGR01090; apt; 1.
```

```
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage.
SQ SEQUENCE 182 AA; 19102 MW; 295D6B4C0C218CD4 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AVRAGL 52
Db 75 AVRAGL 80

RESULT 30
REGA_RHOSU STANDARD; PRT; 183 AA.
AC O82868;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosynthetic apparatus regulatory protein regA.
GN REGA.
OS Rhodovulum sulfidophilum (Rhodobacter sulfidophilus).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodovulum.
OX NCBI_TaxID=35806;
RN [1]
RP SEQUENCE FROM N.A.
RA Masuda S.;
RT "Rhodovulum sulfidophilum photosynthetic regulatory genes.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM REGB/REGA. INVOLVED IN TRANSACTIVATING ANAEROBIC EXPRESSION OF THE PHOTOSYNTHETIC APPARATUS. IT IS A TRANSCRIPTIONAL REGULATOR THAT IS RESPONSIBLE FOR ACTIVATING EXPRESSION OF THE PUF, PUFH, AND PUC OPERONS IN RESPONSE TO A DECREASE IN OXYGEN TENSION (BY SIMILARITY).
CC -!- PTM: Phosphorylated by regB (Probable).
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
EMBL; AB010722; BAA31474.1; -.
DR HSSP; P10958; IDBW.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR008329; Res_reg_RegA.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00072; response_reg; 1.
DR PIRSF; PIRSF036385; Res_reg_RegA; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00448; REC; 1.
DR TIGRFAMS; TIGR01199; HTH_fis; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Sensory transduction; Phosphorylation; Transcription regulation; DNA-binding; Activator.
FT DOMAIN 13 127 RESPONSE_REGULATORY.
FT MOD_RES 62 62 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 183 AA; 20219 MW; 394DB79D39AAB23F CRC64;

Query Match 3.4%; Score 6; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 ASARPP 126
Db 51 ASARPP 56
```

```
RESULT 31
Y432_TREPA
ID Y432_TREPA STANDARD; PRT; 183 AA.
AC O83447;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0432.
GN TP0432.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001220; AAC65420.1; -.
DR PIR; E71326; E71326.
DR TIGR; TP0432; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 183 AA; 19537 MW; 75759BE26C369A35 CRC64;
```

```
Query Match 3.4%; Score 6; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 55 PDYALL 60
Db 90 PDYALL 95
```

```
RESULT 32
Y754_METJA
ID Y754_METJA STANDARD; PRT; 185 AA.
AC Q58164;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0754.
GN MJ0754.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
```

```
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67521; AAB98756.1; -.
DR PIR; B64394; B64394.
DR TIGR; MJ0754; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 185 AA; 21765 MW; 2CE5EC9D424895F6 CRC64;
```

```
Query Match 3.4%; Score 6; DB 1; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 148 NIPDPV 153
Db 69 NIPDPV 74
```

```
RESULT 33
PGHD_SHEEP
ID PGHD_SHEEP STANDARD; PRT; 191 AA.
AC Q9XSM0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Prostaglandin-H2 D-isomerase precursor (EC 5.3.99.2) (Lipocalin-type
DE prostaglandin-D synthase) (Glutathione-independent PGD synthetase)
DE (Prostaglandin D2 synthase) (PGD2 synthase) (PGDS2).
GN PTGDS.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=21663383; PubMed=11804963;
RA Fouchecourt S., Charpigny G., Renaud P., Dumont P., Dacheux J.-L.;
RT "Mammalian lipocalin-type prostaglandin D2 synthase in the fluids of
RT the male genital tract: putative biochemical and physiological
RT functions.";
RL Biol. Reprod. 66:458-467(2002).
RN [2]
RP SEQUENCE OF 29-45; 63-68 AND 100-107, AND TISSUE SPECIFICITY.
RX MEDLINE=99150178; PubMed=10026099;
RA Fouchecourt S., Dacheux F., Dacheux J.-L.;
RT "Glutathione-independent prostaglandin D2 synthase in ram and stallion
RT epididymal fluids: origin and regulation.";
RL Biol. Reprod. 60:558-566(1999).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=96177373; PubMed=8599604;
RA Giacomelli S., Leone M.G., Grima J., Silvestrini B., Cheng C.Y.;
RT "Astrocytes synthesize and secrete prostaglandin D synthetase in
RT vitro.";
RL Biochim. Biophys. Acta 1310:269-276(1996).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=22380479; PubMed=12493710;
RA Fouchecourt S., Castella S., Dacheux F., Dacheux J.-L.;
RT "Prostaglandin d(2) synthase secreted in the caput epididymidis
```





```
RESULT 35
RL9B YEAST
ID RL9B YEAST STANDARD; PRT; 191 AA.
AC P51401;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L9-B (L8) (YL11) (RP25) .
GN RPL9B OR YNL067W OR N2406 OR YNL2406W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96267764; PubMed=8701611;
RA Poehlmann R., Philippsen P.;
RT "Sequencing a cosmid clone of Saccharomyces cerevisiae chromosome XIV
RT reveals 12 new open reading frames (ORFs) and an ancient duplication
RT of six ORFs.";
RL Yeast 12:391-402(1996) .
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1676;
RX MEDLINE=96021608; PubMed=8533472;
RA Bergez P., Daignon F., Crouzet M.;
RT "The sequence of a 44 420 bp fragment located on the left arm of
RT chromosome XIV from Saccharomyces cerevisiae.";
RL Yeast 11:967-974(1995) .
RN [3]
RP ERRATUM.
RX MEDLINE=97060022; PubMed=8904343;
RA Bergez P., Daignon F., Crouzet M.;
RL Yeast 12:297-297(1996) .
CC -!- MISCELLANEOUS: THERE ARE TWO GENES FOR L9 IN YEAST.
CC -!- SIMILARITY: Belongs to the L6P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86470; CAA60195.1; -.
DR EMBL; U12141; AAA99644.1; -.
DR EMBL; Z71343; CAA95940.1; -.
DR PIR; S53915; S53915.
DR HSSP; P02391; 1RL6.
DR GermOnline; 143073; -.
DR SGD; S0005011; RPL9B.
DR InterPro; IPR000702; Ribosomal L6.
DR InterPro; IPR002359; Ribosomal_L6_2.
DR Pfam; PF00347; Ribosomal_L6; 2.
DR PROSITE; PS00700; RIBOSOMAL_L6_2; 1.
KW Ribosomal protein; Multigene family.
SQ SEQUENCE 191 AA; 21657 MW; 27DC5CE8B4B3A4D7 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AALRTV 39
Db 66 AALRTV 71

RESULT 36
AEQ1_AEQVI
```

```
ID AEQ1_AEQVI STANDARD; PRT; 196 AA.
AC P07164;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aequorin 1 precursor.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87185437; PubMed=2882777;
RA Prasher D.C., McCann R.O., Longiaru M., Cormier M.J.;
RT "Sequence comparisons of complementary DNAs encoding aequorin
RT isotypes.";
RL Biochemistry 26:1326-1332(1987) .
RN [2]
RP SEQUENCE OF 8-196.
RX MEDLINE=86077721; PubMed=2866797;
RA Charbonneau H., Walsh K.A., McCann R.O., Prendergast F.G.,
RA Cormier M.J., Vanaman T.C.;
RT "Amino acid sequence of the calcium-dependent photoprotein aequorin.";
RL Biochemistry 24:6762-6771(1985) .
RN [3]
RP MUTAGENESIS OF PRO-196.
RX MEDLINE=92111761; PubMed=1765170;
RA Nomura M., Inouye S., Ohmiya Y., Tsuji F.I.;
RT "A C-terminal proline is required for bioluminescence of the Ca(2+) -
RT binding photoprotein, aequorin.";
RL FEBS Lett. 295:63-66(1991) .
RN [4]
RP DISULFIDE BOND.
RX MEDLINE=94009705; PubMed=8405461;
RA Ohmiya Y., Kurono S., Ohashi M., Fagan T.F., Tsuji F.I.;
RT "Mass spectrometric evidence for a disulfide bond in aequorin
RT regeneration.";
RL FEBS Lett. 332:226-228(1993) .
CC -!- FUNCTION: Ca(2+)-dependent bioluminescence photoprotein. Displays
CC an emission peak at 470 nm (blue light). Trace amounts of calcium
CC ion trigger the intramolecular oxidation of the chromophore,
CC coelenterazine into coelenteramide and CO(2) with the
CC concomitant emission of light.
CC -!- PTM: The reduction of the disulfide bond is necessary to
CC regenerate aequorin from apoaequorin.
CC -!- SIMILARITY: BELONGS TO THE EF-HAND SUPERFAMILY. AEQUORIN FAMILY.
CC -!- SIMILARITY: Contains 3 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16103; AAA27716.1; -.
DR PIR; A26623; A26623.
DR HSSP; P02592; 1EJ3.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 3.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00054; EFh; 3.
DR PROSITE; PS00018; EF_HAND; 3.
KW Photoprotein; Calcium-binding; Luminescence; Repeat.
FT PROPEP 1 7
FT CHAIN 8 196 AEQUORIN 1.
FT SITE 47 57 MAY INTERACT WITH THE CHROMOPHORE.
FT SITE 62 72 MAY INTERACT WITH THE CHROMOPHORE.
FT SITE 107 117 MAY INTERACT WITH THE CHROMOPHORE.
FT CA_BIND 31 42 EF-HAND 1 (BY SIMILARITY).
FT DOMAIN 72 88 ANCESTRAL CALCIUM SITE 2.
FT CA_BIND 124 135 EF-HAND 3 (BY SIMILARITY).
```



```
FT  CA BIND      160      171      EF-HAND 4 (BY SIMILARITY).
FT  DISULFID      152      159
SQ  SEQUENCE      196 AA; 22514 MW; 9AA5B636288A5B8F CRC64;

Query Match      3.4%; Score 6; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GDALFD 77
    |||||
Db 116 GDALFD 121

RESULT 37
AEQ2 AEQVI
ID AEQ2_AEQVI STANDARD; PRT; 196 AA.
AC P02592;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aequorin 2 precursor.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85216460; PubMed=3858813;
RA Inouye S., Noguchi M., Sakaki Y., Takagi Y., Miyata T., Iwanaga S.,
RA Miyata T., Tsuji F.I.;
RT "Cloning and sequence analysis of cDNA for the luminescent protein
RT aequorin.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3154-3158(1985).
RN [2]
RP SEQUENCE OF 9-185 FROM N.A. (AEQUORIN 2 AND 3).
RX MEDLINE=87185437; PubMed=2882777;
RA Prasher D.C., McCann R.O., Longiaru M., Cormier M.J.;
RT "Sequence comparisons of complementary DNAs encoding aequorin
RT isotypes.";
RL Biochemistry 26:1326-1332(1987).
RN [3]
RP SEQUENCE OF 8-196.
RX MEDLINE=86077721; PubMed=2866797;
RA Charbonneau H., Walsh K.A., McCann R.O., Prendergast F.G.,
RA Cormier M.J., Vanaman T.C.;
RT "Amino acid sequence of the calcium-dependent photoprotein aequorin.";
RL Biochemistry 24:6762-6771(1985).
RN [4]
RP MUTAGENESIS.
RA Tsuji F.I., Inouye S., Goto T., Sakaki Y.;
RT "Site-specific mutagenesis of the calcium-binding photoprotein
RT aequorin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8107-8111(1986).
RN [5]
RP MUTAGENESIS OF PRO-196.
RX MEDLINE=92111761; PubMed=1765170;
RA Nomura M., Inouye S., Ohmiya Y., Tsuji F.I.;
RT "A C-terminal proline is required for bioluminescence of the Ca(2+)-
RT binding photoprotein, aequorin.";
RL FEBS Lett. 295:63-66(1991).
RN [6]
RP DISULFIDE BOND.
RX MEDLINE=94009705; PubMed=8405461;
RA Ohmiya Y., Kurono S., Ohashi M., Fagan T.F., Tsuji F.I.;
RT "Mass spectrometric evidence for a disulfide bond in aequorin
RT regeneration.";
RL FEBS Lett. 332:226-228(1993).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=20289815; PubMed=10830969;
RA Head J.F., Inouye S., Teranishi K., Shimomura O.;
RT "The crystal structure of the photoprotein aequorin at 2.3-A
RT resolution.";
```

Nature 405:372-376(2000).

-!- FUNCTION: Ca(2+)-dependent bioluminescence photoprotein. Displays an emission peak at 470 nm (blue light). Trace amounts of calcium ion trigger the intramolecular oxidation of the chromophore, coelenterazine into coelenteramide and CO(2) with the concomitant emission of light.

-!- PTM: The reduction of the disulfide bond is necessary to regenerate aequorin from apoaequorin.

-!- SIMILARITY: BELONGS TO THE EF-HAND SUPERFAMILY. AEQUORIN FAMILY.

-!- SIMILARITY: Contains 3 EF-hand calcium-binding domains.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; L29571; AAA27720.1; -.  
EMBL; M16104; AAA27717.1; -.  
EMBL; M16105; AAA27718.1; -.  
EMBL; M11394; AAA27719.1; -.  
PIR; A03020; AQJFNV.  
PDB; 1EU3; 31-MAY-00.  
InterPro; IPR002048; EF-hand.  
Pfam; PF00036; efhand; 3.  
ProDom; PD000012; EF-hand; 1.  
SMART; SM00054; EFh; 2.  
PROSITE; PS00018; EF\_HAND; 3.  
KW Photoprotein; Calcium-binding; Luminescence; Repeat; 3D-structure.  
FT PROPEP 1 7  
FT CHAIN 8 196 AEQUORIN 2.  
FT SITE 47 57 MAY INTERACT WITH THE CHROMOPHORE.  
FT SITE 62 72 MAY INTERACT WITH THE CHROMOPHORE.  
FT SITE 107 117 MAY INTERACT WITH THE CHROMOPHORE.  
FT CA BIND 31 42 EF-HAND 1 (BY SIMILARITY).  
FT DOMAIN 72 88 ANCESTRAL CALCIUM SITE 2.  
FT CA BIND 124 135 EF-HAND 3 (BY SIMILARITY).  
FT CA BIND 160 171 EF-HAND 4 (BY SIMILARITY).  
FT DISULFID 152 159  
FT SITE 196 196  
FT VARIANT 70 71 REQUIRED FOR BIOLUMINESCENCE.  
FT VARIANT 164 164 EA -> N (IN AEQUORIN 3).  
FT MUTAGEN 36 36 S -> R: 100% ACTIVITY LOSS.  
FT MUTAGEN 129 129 G->R: 51% ACTIVITY LOSS.  
FT MUTAGEN 165 165 G->R: NO ACTIVITY LOSS.  
FT MUTAGEN 65 65 H->P: 100% ACTIVITY LOSS.  
FT MUTAGEN 152 152 C->S: 33% ACTIVITY LOSS.  
FT MUTAGEN 152 152 C->R: 52% ACTIVITY LOSS.  
FT MUTAGEN 159 159 C->S: 41% ACTIVITY LOSS.  
FT MUTAGEN 187 187 C->S: 71% ACTIVITY LOSS.  
FT CONFLICT 37 37 K -> R (IN REF. 2).  
FT TURN 14 15  
FT HELIX 17 30  
FT TURN 32 33  
FT STRAND 37 38  
FT HELIX 40 53  
FT TURN 54 54  
FT HELIX 59 75  
FT TURN 76 77  
FT TURN 80 81  
FT STRAND 84 85  
FT HELIX 86 105  
FT TURN 106 107  
FT HELIX 111 123  
FT TURN 125 126  
FT STRAND 130 131  
FT HELIX 133 143  
FT TURN 144 144  
FT HELIX 149 159  
FT TURN 160 160  
FT TURN 163 164

FT STRAND 167 168  
FT HELIX 169 180  
FT TURN 181 182  
FT HELIX 185 187  
FT TURN 188 194  
SQ SEQUENCE 196 AA; 22285 MW; 532DC7A9D29BA80C CRC64;

Query Match 3.4%; Score 6; DB 1; Length 196;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GDALEFD 77  
| | | | |  
Db 116 GDALEFD 121

RESULT 38

MYTR MITCE  
ID MYTR MITCE STANDARD; PRT; 198 AA.

AC P39047;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Mitrocomin precursor.  
GN ML17.

OS Mitrocoma cellullaria (Halistaurea mitrocoma).  
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;  
OC Mitrocomidae; Mitrocoma.  
OX NCBI\_TaxID=31874;  
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94039781; PubMed=8224198;

RA Pagan T.F., Ohmiya Y., Blinks J.R., Inouye S., Tsuji F.I.;

RT "Cloning, expression and sequence analysis of cDNA for the Ca(2+)-  
binding photoprotein, mitrocomin.";  
RL FEBS Lett. 333:301-305(1993).

CC -!- FUNCTION: CA(++)-DEPENDENT BIOLUMINESCENCE PHOTOPROTEIN. DISPLAYS  
CC AN EMISSION PEAK AT 470 NM (BLUE LIGHT). TRACE AMOUNTS OF CALCIUM  
CC ION TRIGGER THE INTRAMOLECULAR OXIDATION OF THE CHROMOPHORE,  
CC COELENTERAZINE INTO COELENTERAMIDE AND CO(2) WITH THE  
CC CONCOMITANT EMISSION OF LIGHT.  
CC -!- SIMILARITY: BELONGS TO THE EF-HAND SUPERFAMILY. AEQUORIN FAMILY.  
CC -!- SIMILARITY: Contains 3 EF-hand calcium-binding domains.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L31623; AAA29298.1; -.

DR PIR; S39022; S39022.

DR HSSP; P02592; 1EJ3.

DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00036; ehand; 3.

DR ProDom; PD000012; EF-hand; 1.

DR SMART; SM00054; EFh; 3.

DR PROSITE; PS00018; EF HAND; 3.

KW Photoprotein; Calcium-binding; Luminescence; Repeat.

FT PROPEP 1 8 POTENTIAL.

FT CHAIN 9 198 MITROCOMIN.

FT CA BIND 32 43 EF-HAND 1 (POTENTIAL).

FT DOMAIN 73 89 ANCESTRAL CALCIUM SITE 2.

FT CA BIND 125 136 EF-HAND 3 (POTENTIAL).

FT CA\_BIND 161 172 EF-HAND 4 (POTENTIAL).

FT DISULFID 153 160 BY SIMILARITY.

SQ SEQUENCE 198 AA; 22714 MW; 8F6307EF0966F670 CRC64;

Query Match

Best Local Similarity 100.0%; Score 6; DB 1; Length 198;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GDALEFD 77  
| | | | |  
Db 117 GDALEFD 122

RESULT 39

HS27 HUMAN

ID HS27 HUMAN STANDARD; PRT; 205 AA.

AC P04792; Q9UC31;

DT 13-AUG-1987 (Rel. 05, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Heat shock 27 kDa protein (HSP 27) (Stress-responsive protein 27)

DE (SRP27) (Estrogen-regulated 24 kDa protein) (28 kDa heat shock

DE protein).

GN HSPB1 OR HSP27.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86232547; PubMed=3714473;

RA Hickey E., Brandon S.E., Potter R., Stein G., Stein J., Weber L.A.;

RT "Sequence and organization of genes encoding the human 27 kDa heat

RT shock protein.";

RL Nucleic Acids Res. 14:4127-4145(1986).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=91057161; PubMed=2243808;

RA Carper S.W., Rocheleau T.A., Storm F.K.;

RT "cDNA sequence of a human heat shock protein HSP27.";

RL Nucleic Acids Res. 18:6457-6457(1990).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Breast carcinoma;

RA Briolay J., Chareyron P., Mehlen P., Arrigo A.;

RT "Identification of a new cDNA sequence from human breast carcinoma

RT cells encoding the 28kDa heat shock protein.";

RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,

RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RA Kozlowicz A., Bauer C., Ames M., Godfrey J.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND SUBUNIT.

RC TISSUE=Cervical carcinoma;

RX MEDLINE=20241836; PubMed=10777697;

RA Hino M., Kurogi K., Okubo M.-A., Murata-Hori M., Hosoya H.;

RT "Small heat shock protein 27 (Hsp27) associates with

RT tubulin/microtubules in HeLa cells.";

RL Biochem. Biophys. Res. Commun. 271:164-169(2000).

RN [7]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung, and Pancreas;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collings F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,



RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE OF 109-205 FROM N.A.  
RC TISSUE=Breast carcinoma;  
RX MEDLINE=89303813; PubMed=2743305;  
RA Fuqua S.A.W., Blum-Salingaros M., McGuire W.L.;  
RT "Induction of the estrogen-regulated '24K' protein by heat shock.";  
RL Cancer Res. 49:4126-4129(1989).  
RN [9]  
RP SEQUENCE OF 122-205 FROM N.A.  
RX MEDLINE=92107919; PubMed=1763035;  
RA Mendelsohn M.E., Zhu Y., O'Neill S.;  
RT "The 29-kDa proteins phosphorylated in thrombin-activated human  
RT platelets are forms of the estrogen receptor-related 27-kDa heat  
RT shock protein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:11212-11216(1991).  
RN [10]  
RP SEQUENCE OF 13-20; 38-46; 97-110; 141-154 AND 172-186, AND  
RP PHOSPHORYLATION.  
RC TISSUE=Breast carcinoma;  
RX MEDLINE=93315498; PubMed=8325890;  
RA Faucher C., Capdevielle J., Canal I., Ferrara P., Mazarguil H.,  
RA McGuire W.L., Darbon J.-M.;  
RT "The 28-kDa protein whose phosphorylation is induced by protein kinase  
RT C activators in MCF-7 cells belongs to the family of low molecular  
RT mass heat shock proteins and is the estrogen-regulated 24-kDa  
RT protein.";  
RL J. Biol. Chem. 268:15168-15173(1993).  
CC -!- FUNCTION: INVOLVED IN STRESS RESISTANCE AND ACTIN ORGANIZATION.  
CC -!- SUBUNIT: Associates with alpha- and beta-tubulin and microtubules.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic in interphase cells. Colocalizes  
CC with mitotic spindles in mitotic cells. Translocates to the  
CC nucleus during heat shock.  
CC -!- INDUCTION: EXPRESSED IN RESPONSE TO ENVIRONMENTAL STRESSES SUCH AS  
CC HEAT SHOCK, OR ESTROGEN STIMULATION IN MCF-7 CELLS.  
CC -!- PTM: Phosphorylated in MCF-7 cells on exposure to protein kinase C  
CC activators and heat shock.  
CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)  
CC family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; L39370; AAA62175.1; -;  
DR EMBL; X54079; CAA38016.1; -;  
DR EMBL; Z23090; CAA80636.1; -;  
DR EMBL; U90906; AAB51056.1; -;  
DR EMBL; AC006388; -; NOT ANNOTATED\_CDS.  
DR EMBL; AB020027; BAB17232.1; -;  
DR EMBL; BC000510; AAH00510.1; -;  
DR EMBL; BC012768; AAH12768.1; -;  
DR EMBL; X16477; CAA34498.1; -;  
DR EMBL; S74571; AAB20722.1; -;  
DR PIR; S12102; HHHU27.  
DR SWISS-2DPAGE; P04792; HUMAN.  
DR Aarhus/Ghent-2DPAGE; 4110; IEF.  
DR Aarhus/Ghent-2DPAGE; 5102; IEF.  
DR Aarhus/Ghent-2DPAGE; 6104; IEF.

DR HSC-2DPAGE; P04792; HUMAN.  
DR Genew; HGNC:5246; HSPB1.  
DR MIM; 602195; -;  
DR GO; GO:0005737; C:cytoplasm; TAS.  
DR GO; GO:0003773; F:heat shock protein activity; TAS.  
DR GO; GO:0006446; P:regulation of translational initiation; TAS.  
DR InterPro; IPR001436; Crystallin\_alpha.  
DR InterPro; IPR002068; HSP20.  
DR InterPro; IPR008978; HSP20\_chap.  
DR Pfam; PF00011; HSP20; 1.  
DR PRINTS; PRO0299; ACRYSTALLIN.  
DR PROSITE; PS01031; HSP20; 1.  
KW Heat shock; Phosphorylation.  
FT MOD\_RES 15 15  
FT PHOSPHORYLATION (BY PKC AND PKA)  
FT (BY SIMILARITY).  
FT MOD\_RES 82 82  
FT PHOSPHORYLATION (BY PKC AND PKA)  
FT (BY SIMILARITY).  
FT CONFLICT 194 205  
FT PEAKSDETAAK -> RSKIR (IN REF. 1, 8 AND  
FT 9).  
SQ SEQUENCE 205 AA; 22782 MW; 1B4DC44A6F6606D5 CRC64;  
  
Query Match 3.4%; Score 6; DB 1; Length 205;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 108 RPLPPA 113  
Db 56 RPLPPA 61  
|||||  
  
RESULT 40  
TAL\_BACHD STANDARD; PRT; 212 AA.  
ID TAL\_BACHD  
AC Q9K6E4; AC  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable transaldolase (EC 2.2.1.2).  
GN TAL OR BH3785.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
CC -!- FUNCTION: Transaldolase is important for the balance of  
CC metabolites in the pentose-phosphate pathway (By similarity).  
CC -!- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde  
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.  
CC -!- PATHWAY: Pentose phosphate pathway; nonoxidative part.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- SIMILARITY: Belongs to the transaldolase family. Subfamily 3B.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AP001520; BAB07504.1; -;  
DR PIR; A84123; A84123.  
DR HSSP; P30148; IUCW.  
DR HAMAP; MF\_00494; -; 1.  
DR InterPro; IPR001585; Transaldolase.

DR InterPro; IPR004731; Transaldolase\_C.  
DR Pfam; PF00923; Transaldolase; 1.  
DR TIGRFAMs; TIGR00875; talC; 1.  
DR PROSITE; PS01054; TRANSALDOLASE\_1; 1.  
DR PROSITE; PS00958; TRANSALDOLASE\_2; FALSE NEG.  
KW Transferrase; Pentose shunt; Complete proteome.  
FT ACT SITE 83 BY SIMILARITY.  
SQ SEQUENCE 212 AA; 22981 MW; C84EB7FD7C746BD9 CRC64;

Query Match 3.4%; Score 6; DB 1; Length 212;  
Best Local Similarity 100.0%; Pred.No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 TEAALR 37  
| | | | |  
Db 172 TEAALR 177

Search completed: May 18, 2004, 16:20:49  
Job time : 42 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 16:21:55 ; Search time 43 Seconds  
(without alignments)  
1125.988 Million cell updates/sec

Title: US-10-068-956-2  
Perfect score: 174  
Sequence: 1 RGHVVGVAHTLGHNSRGFGV.....SAYAASAQPTQACPPFPSS 174

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1145568 seqs, 278261457 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	174	100.0	174	15	US-10-068-956-2
2	174	100.0	634	14	US-10-239-663-63
3	174	100.0	634	15	US-10-094-749-2263
4	87	50.0	576	14	US-10-239-663-64
5	87	50.0	576	16	US-10-467-248-9
6	10	5.7	530	15	US-10-068-956-4
7	8	4.6	526	12	US-10-206-576-84
8	8	4.6	546	12	US-10-206-576-82
9	7	4.0	13	9	US-09-938-315-93
10	7	4.0	13	14	US-10-161-791-93
11	7	4.0	31	9	US-09-938-315-60
12	7	4.0	31	14	US-10-161-791-60
13	7	4.0	44	9	US-09-864-761-40010
14	7	4.0	93	14	US-10-029-386-30230
15	7	4.0	103	12	US-10-425-114-69630
					Sequence 2, Appli
					Sequence 63, Appl
					Sequence 2263, Ap
					Sequence 64, Appl
					Sequence 9, Appli
					Sequence 4, Appli
					Sequence 84, Appl
					Sequence 82, Appl
					Sequence 93, Appl
					Sequence 60, Appl
					Sequence 40010, A
					Sequence 30230, A
					Sequence 69630, A

16	7	4.0	127	9	US-09-815-242-10432	Sequence 10432, A
17	7	4.0	127	14	US-10-156-761-7920	Sequence 7920, Ap
18	7	4.0	135	12	US-10-424-599-192521	Sequence 192521,
19	7	4.0	167	12	US-10-424-599-190453	Sequence 190453,
20	7	4.0	182	12	US-10-424-599-221448	Sequence 221448,
21	7	4.0	214	12	US-10-425-114-43280	Sequence 43280, A
22	7	4.0	233	12	US-10-282-122A-49294	Sequence 49294, A
23	7	4.0	233	12	US-10-425-114-48990	Sequence 48990, A
24	7	4.0	233	14	US-10-103-313-562	Sequence 562, App
25	7	4.0	236	15	US-10-369-493-20167	Sequence 20167, A
26	7	4.0	314	12	US-10-424-599-231037	Sequence 231037,
27	7	4.0	361	12	US-10-412-699B-1020	Sequence 1020, Ap
28	7	4.0	361	15	US-10-374-780A-525	Sequence 525, App
29	7	4.0	377	15	US-10-210-130-118	Sequence 118, App
30	7	4.0	385	14	US-10-287-274-393	Sequence 393, App
31	7	4.0	387	14	US-10-310-002-13	Sequence 13, Appl
32	7	4.0	428	12	US-10-424-599-144309	Sequence 144309,
33	7	4.0	432	14	US-10-220-380-4	Sequence 4, Appli
34	7	4.0	440	12	US-10-425-114-43085	Sequence 43085, A
35	7	4.0	452	12	US-10-425-114-59071	Sequence 59071, A
36	7	4.0	495	12	US-10-424-599-234625	Sequence 234625,
37	7	4.0	500	12	US-10-425-114-59944	Sequence 59944, A
38	7	4.0	502	12	US-10-308-128-6	Sequence 6, Appli
39	7	4.0	505	15	US-10-266-829-112	Sequence 112, App
40	7	4.0	509	11	US-09-833-245-321	Sequence 321, App
41	7	4.0	509	11	US-09-833-245-2021	Sequence 2021, Ap
42	7	4.0	509	15	US-10-266-829-64	Sequence 64, Appl
43	7	4.0	529	12	US-10-382-248-14	Sequence 14, Appl
44	7	4.0	534	12	US-10-377-097-96	Sequence 96, Appl
45	7	4.0	554	11	US-09-833-245-2020	Sequence 2020, Ap
46	7	4.0	554	15	US-10-266-829-89	Sequence 89, Appl
47	7	4.0	652	15	US-10-369-493-18004	Sequence 18004, A
48	7	4.0	724	15	US-10-369-493-3618	Sequence 3618, Ap
49	7	4.0	735	15	US-10-369-493-2500	Sequence 2500, Ap
50	7	4.0	844	14	US-10-156-761-7650	Sequence 7650, Ap
51	7	4.0	950	9	US-09-823-356-9	Sequence 9, Appli
52	7	4.0	1008	14	US-10-310-002-11	Sequence 11, Appl
53	7	4.0	1047	14	US-10-310-002-12	Sequence 12, Appl
54	7	4.0	1054	14	US-10-310-002-47	Sequence 47, Appl
55	7	4.0	1288	15	US-10-115-479-80	Sequence 80, Appl
56	7	4.0	1408	15	US-10-115-479-82	Sequence 82, Appl
57	7	4.0	1463	14	US-10-310-002-48	Sequence 48, Appl
58	6	3.4	13	9	US-09-938-315-82	Sequence 82, Appl
59	6	3.4	13	9	US-09-938-315-85	Sequence 85, Appl
60	6	3.4	13	9	US-09-938-315-88	Sequence 88, Appl
61	6	3.4	13	9	US-09-938-315-90	Sequence 90, Appl
62	6	3.4	13	14	US-10-161-791-82	Sequence 82, Appl
63	6	3.4	13	14	US-10-161-791-85	Sequence 85, Appl
64	6	3.4	13	14	US-10-161-791-88	Sequence 88, Appl
65	6	3.4	13	15	US-10-283-940-84	Sequence 84, Appl
66	6	3.4	14	10	US-09-933-767-365	Sequence 365, App
67	6	3.4	14	12	US-10-004-860-365	Sequence 365, App
68	6	3.4	14	14	US-10-023-282-365	Sequence 365, App
69	6	3.4	15	14	US-10-161-791-390	Sequence 390, App
70	6	3.4	16	14	US-10-161-791-411	Sequence 411, App
71	6	3.4	17	14	US-10-161-791-370	Sequence 370, App
72	6	3.4	25	10	US-09-820-649-321	Sequence 321, App
73	6	3.4	25	14	US-10-160-162-321	Sequence 321, App
74	6	3.4	31	9	US-09-938-315-53	Sequence 53, Appl
75	6	3.4	31	9	US-09-938-315-56	Sequence 56, Appl
76	6	3.4	31	14	US-10-161-791-53	Sequence 53, Appl
77	6	3.4	31	14	US-10-161-791-56	Sequence 56, Appl
78	6	3.4	39	12	US-10-424-599-228561	Sequence 228561,
79	6	3.4	41	9	US-09-864-761-44546	Sequence 44546, A
80	6	3.4	43	9	US-09-764-877-1787	Sequence 1787, Ap
81	6	3.4	43	15	US-10-242-515-1787	Sequence 1787, Ap
82	6	3.4	44	12	US-10-424-599-233580	Sequence 233580,
83	6	3.4	50	9	US-09-938-315-50	Sequence 50, Appl
84	6	3.4	50	14	US-10-161-791-50	Sequence 50, Appl
85	6	3.4	52	12	US-10-424-599-232325	Sequence 232325,
86	6	3.4	54	9	US-09-764-869-884	Sequence 884, App
87	6	3.4	54	12	US-10-424-599-243173	Sequence 243173,
88	6	3.4	54	14	US-10-091-504-884	Sequence 884, App

89	6	3.4	54	15	US-10-227-577-884	Sequence 884, App	162	6	3.4	119	12	US-10-424-599-226945	Sequence 226945,
90	6	3.4	55	12	US-10-424-599-247279	Sequence 247279,	163	6	3.4	120	10	US-09-809-391-426	Sequence 426, App
91	6	3.4	58	12	US-10-424-599-199452	Sequence 199452,	164	6	3.4	120	10	US-09-882-171-426	Sequence 426, App
92	6	3.4	60	12	US-10-424-599-223589	Sequence 223589,	165	6	3.4	120	12	US-10-164-861-426	Sequence 426, App
93	6	3.4	61	8	US-08-424-550B-468	Sequence 468, App	166	6	3.4	121	9	US-09-738-626-5559	Sequence 5559, Ap
94	6	3.4	61	12	US-10-424-599-243937	Sequence 243937,	167	6	3.4	122	12	US-10-424-599-229517	Sequence 229517,
95	6	3.4	63	12	US-10-424-599-154131	Sequence 154131,	168	6	3.4	124	9	US-09-867-550-914	Sequence 914, App
96	6	3.4	64	9	US-09-864-761-44665	Sequence 44665, A	169	6	3.4	126	12	US-10-424-599-236489	Sequence 236489,
97	6	3.4	68	12	US-10-424-599-142959	Sequence 142959,	170	6	3.4	128	12	US-10-424-599-259211	Sequence 259211,
98	6	3.4	68	12	US-10-424-599-219046	Sequence 219046,	171	6	3.4	129	12	US-10-425-114-48253	Sequence 48253, A
99	6	3.4	68	12	US-10-424-599-273761	Sequence 273761,	172	6	3.4	131	12	US-10-282-122A-61089	Sequence 61089, A
100	6	3.4	69	12	US-10-424-599-187080	Sequence 187080,	173	6	3.4	132	15	US-10-104-047-2471	Sequence 2471, Ap
101	6	3.4	71	12	US-10-424-599-179022	Sequence 179022,	174	6	3.4	137	9	US-09-867-550-524	Sequence 524, App
102	6	3.4	71	12	US-10-424-599-274152	Sequence 274152,	175	6	3.4	138	12	US-10-282-122A-63931	Sequence 63931, A
103	6	3.4	71	12	US-10-425-114-39324	Sequence 39324, A	176	6	3.4	139	12	US-10-424-599-260354	Sequence 260354,
104	6	3.4	72	12	US-10-425-114-55452	Sequence 55452, A	177	6	3.4	140	12	US-10-425-114-49181	Sequence 49181, A
105	6	3.4	74	12	US-10-424-599-217329	Sequence 217329,	178	6	3.4	140	15	US-10-104-047-2801	Sequence 2801, Ap
106	6	3.4	75	12	US-10-424-599-196921	Sequence 196921,	179	6	3.4	141	12	US-10-424-599-231330	Sequence 231330,
107	6	3.4	76	12	US-10-424-599-206901	Sequence 206901,	180	6	3.4	141	12	US-10-424-599-255031	Sequence 255031,
108	6	3.4	76	12	US-10-424-599-225910	Sequence 225910,	181	6	3.4	141	12	US-10-424-599-259235	Sequence 259235,
109	6	3.4	76	12	US-10-424-599-246129	Sequence 246129,	182	6	3.4	141	14	US-10-156-761-7981	Sequence 7981, Ap
110	6	3.4	77	12	US-10-424-599-282590	Sequence 282590,	183	6	3.4	142	12	US-10-424-599-188759	Sequence 188759,
111	6	3.4	77	14	US-10-106-698-8414	Sequence 8414, Ap	184	6	3.4	142	12	US-10-296-115-776	Sequence 776, App
112	6	3.4	78	12	US-10-424-599-191848	Sequence 191848,	185	6	3.4	143	9	US-09-897-214-5	Sequence 5, Appli
113	6	3.4	80	11	US-09-864-408A-8520	Sequence 8520, Ap	186	6	3.4	143	9	US-09-893-737-190	Sequence 190, App
114	6	3.4	82	9	US-09-867-550-468	Sequence 468, App	187	6	3.4	144	12	US-10-424-599-272428	Sequence 272428,
115	6	3.4	82	10	US-09-764-872-380	Sequence 380, App	188	6	3.4	145	12	US-10-424-599-229168	Sequence 229168,
116	6	3.4	82	12	US-10-424-599-269918	Sequence 269918,	189	6	3.4	146	14	US-10-062-831-63	Sequence 63, Appl
117	6	3.4	83	9	US-09-903-456-8	Sequence 8, Appli	190	6	3.4	146	14	US-10-062-831-80	Sequence 80, Appl
118	6	3.4	83	14	US-10-156-911-8	Sequence 8, Appli	191	6	3.4	146	14	US-10-062-599-63	Sequence 63, Appl
119	6	3.4	83	14	US-10-408-736-5	Sequence 5, Appli	192	6	3.4	146	14	US-10-062-599-80	Sequence 80, Appl
120	6	3.4	84	9	US-09-726-643-66	Sequence 66, Appl	193	6	3.4	150	12	US-10-282-122A-62008	Sequence 62008, A
121	6	3.4	84	12	US-10-424-599-225606	Sequence 225606,	194	6	3.4	151	12	US-10-425-114-69400	Sequence 69400, A
122	6	3.4	84	13	US-10-042-141-66	Sequence 66, Appl	195	6	3.4	152	12	US-10-424-599-187618	Sequence 187618,
123	6	3.4	85	12	US-09-925-298-770	Sequence 770, App	196	6	3.4	153	11	US-09-864-408A-5606	Sequence 5606, Ap
124	6	3.4	85	14	US-10-102-806-770	Sequence 770, App	197	6	3.4	154	12	US-10-424-599-197181	Sequence 197181,
125	6	3.4	86	12	US-10-424-599-242330	Sequence 242330,	198	6	3.4	154	12	US-10-425-114-48843	Sequence 48843, A
126	6	3.4	87	12	US-10-425-114-71585	Sequence 71585, A	199	6	3.4	155	12	US-10-424-599-266586	Sequence 266586,
127	6	3.4	88	11	US-09-864-408A-7060	Sequence 7060, Ap	200	6	3.4	155	12	US-10-425-114-47186	Sequence 47186, A
128	6	3.4	90	10	US-09-764-872-370	Sequence 370, App	201	6	3.4	157	12	US-10-282-122A-59894	Sequence 59894, A
129	6	3.4	91	10	US-09-894-159-38	Sequence 38, Appl	202	6	3.4	157	12	US-10-424-599-224175	Sequence 224175,
130	6	3.4	91	12	US-10-424-599-152717	Sequence 152717,	203	6	3.4	157	15	US-10-108-260A-4396	Sequence 4396, Ap
131	6	3.4	92	11	US-09-864-408A-8104	Sequence 8104, Ap	204	6	3.4	158	12	US-10-425-114-60238	Sequence 60238, A
132	6	3.4	92	14	US-10-106-698-7181	Sequence 7181, Ap	205	6	3.4	159	9	US-09-764-869-1038	Sequence 1038, Ap
133	6	3.4	93	9	US-09-925-300-1534	Sequence 1534, Ap	206	6	3.4	159	12	US-10-424-599-188316	Sequence 188316,
134	6	3.4	93	10	US-09-820-649-159	Sequence 159, App	207	6	3.4	159	12	US-10-424-599-218381	Sequence 218381,
135	6	3.4	93	14	US-10-160-162-159	Sequence 159, App	208	6	3.4	159	14	US-10-091-504-1038	Sequence 1038, Ap
136	6	3.4	94	12	US-10-424-599-272549	Sequence 272549,	209	6	3.4	159	15	US-10-227-577-1038	Sequence 1038, Ap
137	6	3.4	94	14	US-10-106-698-5724	Sequence 5724, Ap	210	6	3.4	160	9	US-09-925-299-1192	Sequence 1192, Ap
138	6	3.4	96	12	US-10-424-599-147249	Sequence 147249,	211	6	3.4	160	10	US-09-925-299-1192	Sequence 1192, Ap
139	6	3.4	98	10	US-09-746-783-144	Sequence 144, App	212	6	3.4	160	12	US-10-424-599-214704	Sequence 214704,
140	6	3.4	98	15	US-10-138-588-58	Sequence 58, Appl	213	6	3.4	160	12	US-10-425-114-38867	Sequence 38867, A
141	6	3.4	100	12	US-10-424-599-172957	Sequence 172957,	214	6	3.4	160	14	US-10-301-064-10	Sequence 10, Appl
142	6	3.4	100	12	US-10-424-599-269496	Sequence 269496,	215	6	3.4	161	14	US-10-156-761-10103	Sequence 10103, A
143	6	3.4	102	12	US-10-424-599-188506	Sequence 188506,	216	6	3.4	161	14	US-10-029-386-33432	Sequence 33432, A
144	6	3.4	102	12	US-10-424-599-272290	Sequence 272290,	217	6	3.4	162	12	US-10-425-114-56884	Sequence 56884, A
145	6	3.4	104	12	US-10-424-599-233834	Sequence 233834,	218	6	3.4	163	14	US-10-211-962-77	Sequence 77, Appl
146	6	3.4	104	12	US-10-424-599-280140	Sequence 280140,	219	6	3.4	164	12	US-10-424-599-250644	Sequence 250644,
147	6	3.4	105	12	US-10-449-857A-40	Sequence 40, Appl	220	6	3.4	165	12	US-10-424-599-213101	Sequence 213101,
148	6	3.4	106	12	US-10-424-599-209438	Sequence 209438,	221	6	3.4	166	9	US-09-734-329-6	Sequence 6, Appli
149	6	3.4	106	12	US-10-425-114-61878	Sequence 61878, A	222	6	3.4	166	12	US-10-425-114-42144	Sequence 42144, A
150	6	3.4	110	12	US-10-425-114-52366	Sequence 52366, A	223	6	3.4	168	12	US-10-425-114-43358	Sequence 43358, A
151	6	3.4	111	12	US-10-425-114-46903	Sequence 46903, A	224	6	3.4	170	12	US-10-425-114-44820	Sequence 44820, A
152	6	3.4	113	12	US-10-424-599-158001	Sequence 158001,	225	6	3.4	173	12	US-10-282-122A-70974	Sequence 70974, A
153	6	3.4	113	14	US-10-029-386-34171	Sequence 34171, A	226	6	3.4	173	12	US-10-424-599-285008	Sequence 285008,
154	6	3.4	114	15	US-10-369-493-22992	Sequence 22992, A	227	6	3.4	174	9	US-09-731-872-264	Sequence 264, App
155	6	3.4	115	12	US-10-424-599-177267	Sequence 177267,	228	6	3.4	174	10	US-09-876-997-264	Sequence 264, App
156	6	3.4	115	12	US-10-425-114-51580	Sequence 51580, A	229	6	3.4	174	12	US-10-424-599-263710	Sequence 263710,
157	6	3.4	116	12	US-10-424-599-189930	Sequence 189930,	230	6	3.4	174	12	US-10-424-599-278407	Sequence 278407,
158	6	3.4	116	12	US-10-424-599-230737	Sequence 230737,	231	6	3.4	174	15	US-10-115-482-42	Sequence 42, Appl
159	6	3.4	116	15	US-10-084-846A-22	Sequence 22, Appl	232	6	3.4	174	15	US-10-115-482-44	Sequence 44, Appl
160	6	3.4	117	12	US-10-425-114-68076	Sequence 68076, A	233	6	3.4	175	14	US-10-029-386-32557	Sequence 32557, A
161	6	3.4	118	12	US-10-424-599-271704	Sequence 271704,	234	6	3.4	176	9	US-09-734-017A-54	Sequence 54, Appl



235	6	3.4	176	14	US-10-156-761-12927	Sequence 12927, A	308	6	3.4	230	12	US-10-425-114-71819	Sequence 71819, A
236	6	3.4	178	12	US-10-424-599-192572	Sequence 192572,	309	6	3.4	231	12	US-10-424-599-258908	Sequence 258908,
237	6	3.4	178	12	US-10-424-599-192574	Sequence 192574,	310	6	3.4	232	12	US-10-424-599-248662	Sequence 248662,
238	6	3.4	178	12	US-10-425-114-51253	Sequence 51253, A	311	6	3.4	232	14	US-10-156-761-10287	Sequence 10287, A
239	6	3.4	178	14	US-10-193-002-231	Sequence 231, App	312	6	3.4	233	12	US-10-425-114-57682	Sequence 57682, A
240	6	3.4	178	14	US-10-084-843-236	Sequence 236, App	313	6	3.4	234	12	US-10-282-122A-57582	Sequence 57582, A
241	6	3.4	179	12	US-10-425-114-58720	Sequence 58720, A	314	6	3.4	234	12	US-10-424-599-215741	Sequence 215741,
242	6	3.4	180	12	US-10-424-599-202355	Sequence 202355,	315	6	3.4	235	14	US-10-080-170-550	Sequence 550, App
243	6	3.4	180	12	US-10-424-599-228503	Sequence 228503,	316	6	3.4	235	15	US-10-104-047-3463	Sequence 3463, Ap
244	6	3.4	185	10	US-09-791-279-162	Sequence 162, App	317	6	3.4	237	12	US-10-424-599-237185	Sequence 237185,
245	6	3.4	185	12	US-10-424-599-253188	Sequence 253188,	318	6	3.4	237	14	US-10-017-161-954	Sequence 954, App
246	6	3.4	185	12	US-10-425-114-54776	Sequence 54776, A	319	6	3.4	237	15	US-10-104-047-3084	Sequence 3084, Ap
247	6	3.4	186	13	US-10-108-605-121	Sequence 121, App	320	6	3.4	238	9	US-09-764-868-1034	Sequence 1034, Ap
248	6	3.4	188	12	US-10-282-122A-48523	Sequence 48523, A	321	6	3.4	238	16	US-10-389-566-1544	Sequence 1544, Ap
249	6	3.4	188	12	US-10-424-599-263472	Sequence 263472,	322	6	3.4	239	12	US-10-282-122A-67798	Sequence 67798, A
250	6	3.4	189	10	US-09-894-159-28	Sequence 28, Appl	323	6	3.4	239	15	US-10-369-493-7928	Sequence 7928, Ap
251	6	3.4	189	12	US-10-400-630-1	Sequence 1, Appli	324	6	3.4	241	14	US-10-156-761-14265	Sequence 14265, A
252	6	3.4	191	10	US-09-764-891-4152	Sequence 4152, Ap	325	6	3.4	242	15	US-10-289-762-182	Sequence 182, App
253	6	3.4	191	12	US-10-425-114-70073	Sequence 70073, A	326	6	3.4	243	12	US-10-425-114-37315	Sequence 37315, A
254	6	3.4	192	12	US-10-282-122A-45258	Sequence 45258, A	327	6	3.4	244	9	US-09-734-329-5	Sequence 5, Appli
255	6	3.4	192	12	US-10-400-630-2	Sequence 2, Appli	328	6	3.4	245	12	US-10-425-114-53957	Sequence 53957, A
256	6	3.4	194	12	US-10-282-122A-47721	Sequence 47721, A	329	6	3.4	246	9	US-09-738-626-5858	Sequence 5858, Ap
257	6	3.4	196	12	US-10-425-114-48598	Sequence 48598, A	330	6	3.4	246	12	US-10-424-599-265158	Sequence 265158,
258	6	3.4	196	14	US-10-281-013-2	Sequence 2, Appli	331	6	3.4	246	12	US-10-425-114-38255	Sequence 38255, A
259	6	3.4	196	14	US-10-280-911-2	Sequence 2, Appli	332	6	3.4	247	9	US-09-835-996A-27	Sequence 27, Appl
260	6	3.4	196	14	US-10-167-831-95	Sequence 95, Appl	333	6	3.4	247	12	US-10-425-114-66513	Sequence 66513, A
261	6	3.4	198	12	US-10-425-114-42720	Sequence 42720, A	334	6	3.4	247	15	US-10-104-047-2790	Sequence 2790, Ap
262	6	3.4	198	12	US-10-671-403-88	Sequence 88, Appl	335	6	3.4	249	8	US-08-899-112-11	Sequence 11, Appl
263	6	3.4	198	12	US-10-671-419-88	Sequence 88, Appl	336	6	3.4	249	12	US-10-412-699B-122	Sequence 122, App
264	6	3.4	198	12	US-10-670-844-88	Sequence 88, Appl	337	6	3.4	249	12	US-10-225-066A-438	Sequence 438, App
265	6	3.4	198	12	US-10-671-134-88	Sequence 88, Appl	338	6	3.4	249	14	US-10-286-264-148	Sequence 148, App
266	6	3.4	198	12	US-10-673-098-88	Sequence 88, Appl	339	6	3.4	249	15	US-10-369-493-18657	Sequence 18657, A
267	6	3.4	198	14	US-10-219-220-151	Sequence 151, App	340	6	3.4	249	15	US-10-225-067-126	Sequence 126, App
268	6	3.4	198	16	US-10-672-638-88	Sequence 88, Appl	341	6	3.4	249	15	US-10-374-780A-34	Sequence 34, Appl
269	6	3.4	198	16	US-10-673-127-88	Sequence 88, Appl	342	6	3.4	250	12	US-10-425-114-66419	Sequence 66419, A
270	6	3.4	201	12	US-10-425-114-66982	Sequence 66982, A	343	6	3.4	250	15	US-10-369-493-10605	Sequence 10605, A
271	6	3.4	203	12	US-10-282-122A-47383	Sequence 47383, A	344	6	3.4	251	12	US-10-424-599-276797	Sequence 276797,
272	6	3.4	205	14	US-10-153-668-284	Sequence 284, App	345	6	3.4	252	12	US-10-425-114-61954	Sequence 61954, A
273	6	3.4	205	15	US-10-116-275-148	Sequence 148, App	346	6	3.4	252	12	US-10-671-403-180	Sequence 180, App
274	6	3.4	206	12	US-10-424-599-224421	Sequence 224421,	347	6	3.4	252	12	US-10-671-419-180	Sequence 180, App
275	6	3.4	207	12	US-10-425-114-40435	Sequence 40435, A	348	6	3.4	252	12	US-10-670-844-180	Sequence 180, App
276	6	3.4	208	12	US-10-282-122A-47545	Sequence 47545, A	349	6	3.4	252	12	US-10-671-134-180	Sequence 180, App
277	6	3.4	209	12	US-10-424-599-237184	Sequence 237184,	350	6	3.4	252	12	US-10-673-098-180	Sequence 180, App
278	6	3.4	209	12	US-10-425-114-43462	Sequence 43462, A	351	6	3.4	252	15	US-10-369-493-4803	Sequence 4803, Ap
279	6	3.4	209	12	US-10-425-114-48779	Sequence 48779, A	352	6	3.4	252	15	US-10-369-493-7562	Sequence 7562, Ap
280	6	3.4	210	12	US-10-671-403-95	Sequence 95, Appl	353	6	3.4	252	15	US-10-369-493-18557	Sequence 18557, A
281	6	3.4	210	12	US-10-671-419-95	Sequence 95, Appl	354	6	3.4	252	15	US-10-369-493-22893	Sequence 22893, A
282	6	3.4	210	12	US-10-670-844-95	Sequence 95, Appl	355	6	3.4	252	16	US-10-672-638-180	Sequence 180, App
283	6	3.4	210	12	US-10-671-134-95	Sequence 95, Appl	356	6	3.4	252	16	US-10-673-127-180	Sequence 11227, A
284	6	3.4	210	12	US-10-673-098-95	Sequence 95, Appl	357	6	3.4	253	14	US-10-156-761-11227	Sequence 11227, A
285	6	3.4	210	16	US-10-672-638-95	Sequence 95, Appl	358	6	3.4	255	12	US-10-424-599-181403	Sequence 181403,
286	6	3.4	210	16	US-10-673-127-95	Sequence 95, Appl	359	6	3.4	255	15	US-10-369-493-2944	Sequence 2944, Ap
287	6	3.4	211	12	US-10-425-114-54325	Sequence 54325, A	360	6	3.4	255	15	US-10-104-047-3876	Sequence 3876, Ap
288	6	3.4	213	12	US-10-425-114-64548	Sequence 64548, A	361	6	3.4	256	12	US-10-424-599-220803	Sequence 220803,
289	6	3.4	213	14	US-10-156-761-10414	Sequence 10414, A	362	6	3.4	256	12	US-10-425-114-62576	Sequence 62576, A
290	6	3.4	216	15	US-10-369-493-20239	Sequence 20239, A	363	6	3.4	256	15	US-10-369-493-17754	Sequence 17754, A
291	6	3.4	216	15	US-10-041-615-66	Sequence 66, Appl	364	6	3.4	257	12	US-10-282-122A-66732	Sequence 66732, A
292	6	3.4	216	15	US-10-041-615-85	Sequence 85, Appl	365	6	3.4	258	12	US-10-424-599-248969	Sequence 248969,
293	6	3.4	217	15	US-10-369-493-9434	Sequence 9434, Ap	366	6	3.4	258	12	US-10-425-114-47317	Sequence 47317, A
294	6	3.4	217	15	US-10-289-762-656	Sequence 656, App	367	6	3.4	259	12	US-10-424-599-212318	Sequence 212318,
295	6	3.4	218	12	US-10-258-080-9	Sequence 9, Appli	368	6	3.4	259	15	US-10-392-301-34	Sequence 34, Appl
296	6	3.4	218	14	US-10-029-386-32067	Sequence 32067, A	369	6	3.4	260	9	US-09-738-626-4051	Sequence 4051, Ap
297	6	3.4	222	12	US-10-424-599-153456	Sequence 153456,	370	6	3.4	261	9	US-09-925-301-901	Sequence 901, App
298	6	3.4	222	14	US-10-157-031-71	Sequence 71, Appl	371	6	3.4	263	12	US-10-282-122A-49946	Sequence 49946, A
299	6	3.4	222	15	US-10-240-145-161	Sequence 161, App	372	6	3.4	263	14	US-10-156-761-11386	Sequence 11386, A
300	6	3.4	223	15	US-10-374-780A-1897	Sequence 1897, Ap	373	6	3.4	265	12	US-10-424-599-220521	Sequence 220521,
301	6	3.4	225	12	US-10-424-599-174550	Sequence 174550,	374	6	3.4	266	9	US-09-778-963A-2	Sequence 2, Appli
302	6	3.4	227	15	US-10-369-493-8979	Sequence 8979, Ap	375	6	3.4	266	9	US-09-778-963A-4	Sequence 4, Appli
303	6	3.4	228	12	US-10-092-900A-298	Sequence 298, App	376	6	3.4	266	9	US-09-778-963A-5	Sequence 5, Appli
304	6	3.4	228	12	US-10-403-161-64	Sequence 64, Appl	377	6	3.4	266	12	US-10-282-122A-53424	Sequence 53424, A
305	6	3.4	228	15	US-10-369-493-20929	Sequence 20929, A	378	6	3.4	266	12	US-09-918-715-198	Sequence 198, App
306	6	3.4	229	12	US-10-282-122A-55992	Sequence 55992, A	379	6	3.4	266	12	US-09-918-715-293	Sequence 293, App
307	6	3.4	229	12	US-10-425-114-54022	Sequence 54022, A	380	6	3.4	267	12	US-10-425-114-63233	Sequence 63233, A



381	6	3.4	268	9	US-09-738-626-4858	Sequence 4858, Ap	454	6	3.4	280	10	US-09-993-583-319	Sequence 319, App
382	6	3.4	268	12	US-10-115-123-346	Sequence 346, App	455	6	3.4	280	10	US-09-941-992-319	Sequence 319, App
383	6	3.4	268	14	US-10-012-542-346	Sequence 346, App	456	6	3.4	280	10	US-09-992-521-319	Sequence 319, App
384	6	3.4	270	9	US-09-350-874-59	Sequence 59, Appl	457	6	3.4	280	10	US-09-997-333-319	Sequence 319, App
385	6	3.4	270	14	US-10-106-989-59	Sequence 59, Appl	458	6	3.4	280	10	US-09-997-384-319	Sequence 319, App
386	6	3.4	271	9	US-09-864-761-37986	Sequence 37986, A	459	6	3.4	280	10	US-09-998-041-319	Sequence 319, App
387	6	3.4	271	14	US-10-029-386-32518	Sequence 32518, A	460	6	3.4	280	10	US-09-997-585-319	Sequence 319, App
388	6	3.4	272	12	US-10-282-122A-66764	Sequence 66764, A	461	6	3.4	280	10	US-09-997-614-319	Sequence 319, App
389	6	3.4	275	9	US-09-738-626-4402	Sequence 4402, Ap	462	6	3.4	280	10	US-09-989-862-319	Sequence 319, App
390	6	3.4	275	12	US-10-282-122A-49464	Sequence 49464, A	463	6	3.4	280	10	US-09-997-529-319	Sequence 319, App
391	6	3.4	275	12	US-10-412-699B-1033	Sequence 1033, Ap	464	6	3.4	280	10	US-09-989-725-319	Sequence 319, App
392	6	3.4	275	14	US-10-174-209-22	Sequence 22, Appl	465	6	3.4	280	11	US-09-989-733-319	Sequence 319, App
393	6	3.4	275	15	US-10-374-780A-540	Sequence 540, App	466	6	3.4	280	11	US-09-992-643-319	Sequence 319, App
394	6	3.4	277	14	US-10-174-209-20	Sequence 20, Appl	467	6	3.4	280	12	US-10-147-493-458	Sequence 458, App
395	6	3.4	278	12	US-10-425-114-42449	Sequence 42449, A	468	6	3.4	280	12	US-10-206-915-602	Sequence 602, App
396	6	3.4	278	12	US-09-918-715-178	Sequence 178, App	469	6	3.4	280	12	US-10-145-127-458	Sequence 458, App
397	6	3.4	278	14	US-10-080-170-6	Sequence 6, Appli	470	6	3.4	280	12	US-10-160-503-458	Sequence 458, App
398	6	3.4	279	12	US-10-425-114-39970	Sequence 39970, A	471	6	3.4	280	12	US-10-199-670-602	Sequence 602, App
399	6	3.4	280	9	US-09-989-722-319	Sequence 319, App	472	6	3.4	280	12	US-10-143-118-458	Sequence 458, App
400	6	3.4	280	9	US-09-989-723-319	Sequence 319, App	473	6	3.4	280	12	US-10-144-993-458	Sequence 458, App
401	6	3.4	280	9	US-09-989-279-319	Sequence 319, App	474	6	3.4	280	12	US-10-158-787-458	Sequence 458, App
402	6	3.4	280	9	US-09-989-727-319	Sequence 319, App	475	6	3.4	280	12	US-10-201-858-602	Sequence 602, App
403	6	3.4	280	9	US-09-989-731-319	Sequence 319, App	476	6	3.4	280	12	US-10-081-056-166	Sequence 166, App
404	6	3.4	280	9	US-09-989-732-319	Sequence 319, App	477	6	3.4	280	12	US-10-205-890-602	Sequence 602, App
405	6	3.4	280	9	US-09-991-073-319	Sequence 319, App	478	6	3.4	280	12	US-10-208-024-602	Sequence 602, App
406	6	3.4	280	9	US-09-990-442-319	Sequence 319, App	479	6	3.4	280	12	US-10-201-853-602	Sequence 602, App
407	6	3.4	280	9	US-09-991-163-319	Sequence 319, App	480	6	3.4	280	12	US-10-140-024-458	Sequence 458, App
408	6	3.4	280	9	US-09-993-604-319	Sequence 319, App	481	6	3.4	280	12	US-09-989-724-319	Sequence 319, App
409	6	3.4	280	9	US-09-990-456-319	Sequence 319, App	482	6	3.4	280	12	US-09-989-728-319	Sequence 319, App
410	6	3.4	280	9	US-09-835-996A-23	Sequence 23, Appl	483	6	3.4	280	12	US-09-990-441-319	Sequence 319, App
411	6	3.4	280	9	US-09-835-996A-35	Sequence 35, Appl	484	6	3.4	280	12	US-10-140-808-458	Sequence 458, App
412	6	3.4	280	9	US-09-989-721-319	Sequence 319, App	485	6	3.4	280	12	US-10-174-581-602	Sequence 602, App
413	6	3.4	280	9	US-09-992-598-319	Sequence 319, App	486	6	3.4	280	12	US-10-176-483-602	Sequence 602, App
414	6	3.4	280	9	US-09-989-293A-319	Sequence 319, App	487	6	3.4	280	12	US-10-176-749-602	Sequence 602, App
415	6	3.4	280	9	US-09-989-735-319	Sequence 319, App	488	6	3.4	280	12	US-10-176-914-602	Sequence 602, App
416	6	3.4	280	9	US-09-990-444-319	Sequence 319, App	489	6	3.4	280	12	US-10-176-915-602	Sequence 602, App
417	6	3.4	280	9	US-09-991-181-319	Sequence 319, App	490	6	3.4	280	12	US-09-997-857-319	Sequence 319, App
418	6	3.4	280	9	US-09-989-730-319	Sequence 319, App	491	6	3.4	280	12	US-10-176-484-602	Sequence 602, App
419	6	3.4	280	9	US-09-990-436-319	Sequence 319, App	492	6	3.4	280	12	US-10-180-550-602	Sequence 602, App
420	6	3.4	280	9	US-09-993-687-319	Sequence 319, App	493	6	3.4	280	12	US-10-183-014-602	Sequence 602, App
421	6	3.4	280	10	US-09-989-734-319	Sequence 319, App	494	6	3.4	280	12	US-10-187-738-602	Sequence 602, App
422	6	3.4	280	10	US-09-997-653-319	Sequence 319, App	495	6	3.4	280	12	US-10-187-740-602	Sequence 602, App
423	6	3.4	280	10	US-09-993-667-319	Sequence 319, App	496	6	3.4	280	12	US-10-187-883-602	Sequence 602, App
424	6	3.4	280	10	US-09-997-428-319	Sequence 319, App	497	6	3.4	280	12	US-10-194-363-602	Sequence 602, App
425	6	3.4	280	10	US-09-997-666-319	Sequence 319, App	498	6	3.4	280	12	US-10-194-460-602	Sequence 602, App
426	6	3.4	280	10	US-09-990-438-319	Sequence 319, App	499	6	3.4	280	12	US-10-194-463-602	Sequence 602, App
427	6	3.4	280	10	US-09-990-562-319	Sequence 319, App	500	6	3.4	280	12	US-10-194-484-602	Sequence 602, App
428	6	3.4	280	10	US-09-990-711-319	Sequence 319, App	501	6	3.4	280	12	US-10-195-884-602	Sequence 602, App
429	6	3.4	280	10	US-09-989-726-319	Sequence 319, App	502	6	3.4	280	12	US-10-195-896-602	Sequence 602, App
430	6	3.4	280	10	US-09-998-156-319	Sequence 319, App	503	6	3.4	280	12	US-10-196-744-602	Sequence 602, App
431	6	3.4	280	10	US-09-746-783-160	Sequence 160, App	504	6	3.4	280	12	US-10-196-755-602	Sequence 602, App
432	6	3.4	280	10	US-09-990-437-319	Sequence 319, App	505	6	3.4	280	12	US-10-196-757-602	Sequence 602, App
433	6	3.4	280	10	US-09-991-157-319	Sequence 319, App	506	6	3.4	280	12	US-10-197-704-602	Sequence 602, App
434	6	3.4	280	10	US-09-997-514-319	Sequence 319, App	507	6	3.4	280	12	US-10-197-710-602	Sequence 602, App
435	6	3.4	280	10	US-09-997-573-319	Sequence 319, App	508	6	3.4	280	12	US-10-198-758-602	Sequence 602, App
436	6	3.4	280	10	US-09-991-172-319	Sequence 319, App	509	6	3.4	280	12	US-10-198-766-602	Sequence 602, App
437	6	3.4	280	10	US-09-990-726-319	Sequence 319, App	510	6	3.4	280	12	US-10-199-304-602	Sequence 602, App
438	6	3.4	280	10	US-09-997-559-319	Sequence 319, App	511	6	3.4	280	12	US-10-199-309-602	Sequence 602, App
439	6	3.4	280	10	US-09-997-601-319	Sequence 319, App	512	6	3.4	280	12	US-10-199-313-602	Sequence 602, App
440	6	3.4	280	10	US-09-990-443-319	Sequence 319, App	513	6	3.4	280	12	US-10-199-456-602	Sequence 602, App
441	6	3.4	280	10	US-09-991-854-319	Sequence 319, App	514	6	3.4	280	12	US-10-201-329-602	Sequence 602, App
442	6	3.4	280	10	US-09-997-628-319	Sequence 319, App	515	6	3.4	280	12	US-10-202-412-602	Sequence 602, App
443	6	3.4	280	10	US-09-997-683-319	Sequence 319, App	516	6	3.4	280	12	US-10-206-919-602	Sequence 602, App
444	6	3.4	280	10	US-09-989-729A-319	Sequence 319, App	517	6	3.4	280	12	US-10-206-922-602	Sequence 602, App
445	6	3.4	280	10	US-09-997-349-319	Sequence 319, App	518	6	3.4	280	12	US-10-206-924-602	Sequence 602, App
446	6	3.4	280	10	US-09-997-440-319	Sequence 319, App	519	6	3.4	280	12	US-10-206-928-602	Sequence 602, App
447	6	3.4	280	10	US-09-990-440-319	Sequence 319, App	520	6	3.4	280	12	US-10-207-914-602	Sequence 602, App
448	6	3.4	280	10	US-09-993-469-319	Sequence 319, App	521	6	3.4	280	12	US-10-207-921-602	Sequence 602, App
449	6	3.4	280	10	US-09-997-542-319	Sequence 319, App	522	6	3.4	280	12	US-10-207-922-602	Sequence 602, App
450	6	3.4	280	10	US-09-993-748-319	Sequence 319, App	523	6	3.4	280	12	US-10-208-027-602	Sequence 602, App
451	6	3.4	280	10	US-09-990-439-319	Sequence 319, App	524	6	3.4	280	12	US-09-997-641-319	Sequence 319, App
452	6	3.4	280	10	US-09-990-427-319	Sequence 319, App	525	6	3.4	280	12	US-09-991-150-319	Sequence 319, App
453	6	3.4	280	10	US-09-989-328-319	Sequence 319, App	526	6	3.4	280	12	US-10-152-405-458	Sequence 458, App











965	6	3.4	280	14	US-10-121-062-602	Sequence 602, App
966	6	3.4	280	14	US-10-123-215-458	Sequence 458, App
967	6	3.4	280	14	US-10-123-902-458	Sequence 458, App
968	6	3.4	280	14	US-10-123-908-458	Sequence 458, App
969	6	3.4	280	14	US-10-123-909-458	Sequence 458, App
970	6	3.4	280	14	US-10-123-910-458	Sequence 458, App
971	6	3.4	280	14	US-10-124-813-458	Sequence 458, App
972	6	3.4	280	14	US-10-124-817-458	Sequence 458, App
973	6	3.4	280	14	US-10-125-922-458	Sequence 458, App
974	6	3.4	280	14	US-10-125-924-458	Sequence 458, App
975	6	3.4	280	14	US-10-140-860-458	Sequence 458, App
976	6	3.4	280	14	US-10-142-417-458	Sequence 458, App
977	6	3.4	280	14	US-10-147-519-458	Sequence 458, App
978	6	3.4	280	14	US-10-157-782-458	Sequence 458, App
979	6	3.4	280	14	US-10-152-395-458	Sequence 458, App
980	6	3.4	280	14	US-10-125-926A-458	Sequence 458, App
981	6	3.4	280	14	US-10-125-930A-458	Sequence 458, App
982	6	3.4	280	14	US-10-127-831A-458	Sequence 458, App
983	6	3.4	280	14	US-10-127-837A-458	Sequence 458, App
984	6	3.4	280	14	US-10-127-838B-458	Sequence 458, App
985	6	3.4	280	14	US-10-127-842A-458	Sequence 458, App
986	6	3.4	280	14	US-10-127-843A-458	Sequence 458, App
987	6	3.4	280	14	US-10-127-845A-458	Sequence 458, App
988	6	3.4	280	14	US-10-127-846A-458	Sequence 458, App
989	6	3.4	280	14	US-10-127-849A-458	Sequence 458, App
990	6	3.4	280	14	US-10-127-850A-458	Sequence 458, App
991	6	3.4	280	14	US-10-127-851A-458	Sequence 458, App
992	6	3.4	280	14	US-10-128-684A-458	Sequence 458, App
993	6	3.4	280	14	US-10-128-686A-458	Sequence 458, App
994	6	3.4	280	14	US-10-128-690A-458	Sequence 458, App
995	6	3.4	280	14	US-10-128-691A-458	Sequence 458, App
996	6	3.4	280	14	US-10-131-819A-458	Sequence 458, App
997	6	3.4	280	14	US-10-131-829A-458	Sequence 458, App
998	6	3.4	280	14	US-10-131-836A-458	Sequence 458, App
999	6	3.4	280	14	US-10-139-963-458	Sequence 458, App
1000	6	3.4	280	15	US-10-139-963-458	Sequence 458, App

ALIGNMENTS

```
RESULT 1
US-10-068-956-2
; Sequence 2, Application US/10068956
; Publication No. US20030204065A1
; GENERAL INFORMATION:
; APPLICANT: Paul Young et al.
; TITLE OF INVENTION: PGRP-L Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PF513PI
; CURRENT APPLICATION NUMBER: US/10/068,956
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/149,715
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: PCT/US00/22877
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 174
; TYPE: PRT
; ORGANISM: human
US-10-068-956-2

Query Match      100.0%; Score 174; DB 15; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.1e-156;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  RGWHVVGAAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60
      |||||||
Db      1  RGWHVVGAAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60
      |||||||

QY      61 GHRQLVRTDCPGDALFDLLRTPHFTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 120
      |||||||
```

```
Db      61 GHRQLVRTDCPGDALFDLLRTPHFTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 120
QY      121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQPACFPSS 174
      |||||||
Db      121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQPACFPSS 174
      |||||||

RESULT 2
US-10-239-663-63
; Sequence 63, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Konick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,168
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-663-63
```

```
Query Match      100.0%; Score 174; DB 14; Length 634;
Best Local Similarity 100.0%; Pred. No. 3.2e-156;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  RGWHVVGAAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 60
      |||||||
Db      461 RGWHVVGAAHTLGHNSRGFGVAIVGNYTAALPTEAALRTVRDTPSCAVRAGLLRPDYALL 520
      |||||||

QY      61 GHRQLVRTDCPGDALFDLLRTPHFTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 120
      |||||||
Db      521 GHRQLVRTDCPGDALFDLLRTPHFTAVSLRSLHYTARRPSVYTSSTRPLPPACNSCART 580
      |||||||

QY      121 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQPACFPSS 174
      |||||||
Db      581 ASARPPTSRHHVYSGNLGPAFAGHSAGNIPDPVTSAYAASAQPTQPACFPSS 634
      |||||||
```

```
RESULT 3
US-10-094-749-2263
; Sequence 2263, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
```



APPLICANT: SEKI, NAOHIKO  
APPLICANT: YOSHIKAWA, TSUTOMU  
APPLICANT: OTSUKA, MOTOYUKI  
APPLICANT: NAGAHARI, KENJI  
APPLICANT: MASUHO, YASUHIKO  
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
FILE REFERENCE: 084335/0160  
CURRENT APPLICATION NUMBER: US/10/094,749  
CURRENT FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: 60/350,435  
PRIOR FILING DATE: 2002-01-24  
PRIOR APPLICATION NUMBER: JP 2001-328381  
PRIOR FILING DATE: 2001-09-14  
NUMBER OF SEQ ID NOS: 3381  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2263  
LENGTH: 634  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-094-749-2263

Query Match 100.0%; Score 174; DB 15; Length 634;  
Best Local Similarity 100.0%; Pred. No. 3.2e-156;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGHVWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 60  
Db 461 RGHVWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 520  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTFVAVLSLHYTARRPSVYTSSTRPLPPACNSCART 120  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHTFVAVLSLHYTARRPSVYTSSTRPLPPACNSCART 580  
QY 121 ASARPPTSRHRVYSGNLGPAFAGHSAGNIPDPVTSAYAASAPQOTQPCFPSS 174  
Db 581 ASARPPTSRHRVYSGNLGPAFAGHSAGNIPDPVTSAYAASAPQOTQPCFPSS 634

RESULT 4  
US-10-239-663-64  
Sequence 64, Application US/10239663  
Publication No. US20030139572A1  
GENERAL INFORMATION:  
APPLICANT: Agarwal, Pankaj  
APPLICANT: Murdoch, Paul R.  
APPLICANT: Rizvi, Safia, K.  
APPLICANT: Smith, Randall, F.  
APPLICANT: Xiang, Zhaoying  
APPLICANT: Kabnick, Karen  
TITLE OF INVENTION: NOVEL COMPOUNDS  
FILE REFERENCE: GP50018  
CURRENT APPLICATION NUMBER: US/10/239,663  
CURRENT FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: PCT/US01/09226  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: 60/192,158  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 60/192,668  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: 60/200,166  
PRIOR FILING DATE: 2000-04-27  
NUMBER OF SEQ ID NOS: 66  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 64  
LENGTH: 576  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-239-663-64

Query Match 50.0%; Score 87; DB 14; Length 576;  
Best Local Similarity 100.0%; Pred. No. 8e-74;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGHVWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 60  
Db 461 RGHVWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 520  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTF 87  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHTF 547

RESULT 5  
US-10-467-248-9  
Sequence 9, Application US/10467248  
Publication No. US20040086905A1  
GENERAL INFORMATION:  
APPLICANT: DAS, Debopriya; YAO, Monique G.;  
APPLICANT: ARVIZU, Chandra S.; BAUGHN, Mariah R.;  
APPLICANT: LU, Yan; HAFALIA, April J.A.;  
APPLICANT: CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;  
APPLICANT: LU, Dyung Aina M.; YUE, Henry;  
APPLICANT: DING, Li; ELLIOTT, Vicki S.;  
APPLICANT: FORSYTHE, Ian J.; RAMKUMAR, Jayalaxmi;  
APPLICANT: GANDHI, Ameena R.; ISON, Craig H.;  
APPLICANT: WARREN, Bridget A.; TANG, Y. Tom;  
APPLICANT: EMERLING, Brooke M.; HONCHELL, Cynthia D.;  
APPLICANT: LYNE, Michael; BARROSO, Ines  
TITLE OF INVENTION: LIPID-ASSOCIATED MOLECULES  
FILE REFERENCE: PI-0358 USN  
CURRENT APPLICATION NUMBER: US/10/467,248  
CURRENT FILING DATE: 2003-08-06  
PRIOR APPLICATION NUMBER: PCT/US02/03813  
PRIOR FILING DATE: 2002-02-06  
PRIOR APPLICATION NUMBER: US 60/266,910  
PRIOR FILING DATE: 2001-02-06  
PRIOR APPLICATION NUMBER: US 60/276,891  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: US 60/276,855  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: US 60/279,760  
PRIOR FILING DATE: 2001-03-28  
PRIOR APPLICATION NUMBER: US 60/283,818  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: US 60/285,405  
PRIOR FILING DATE: 2001-04-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PERL Program  
SEQ ID NO 9  
LENGTH: 576  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: 6897474CD1  
US-10-467-248-9

Query Match 50.0%; Score 87; DB 16; Length 576;  
Best Local Similarity 100.0%; Pred. No. 8e-74;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGHVWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 60  
Db 461 RGHVWVGAHTLGHNSRGFGVAIVGNVYTAALPTEAALRTVRDTLPSCAVRAGLLRPDYALL 520  
QY 61 GHRQLVRTDCPGDALFDLLRTWPHTF 87  
Db 521 GHRQLVRTDCPGDALFDLLRTWPHTF 547

RESULT 6  
US-10-068-956-4  
Sequence 4, Application US/10068956  
Publication No. US20030204065A1  
GENERAL INFORMATION:  
APPLICANT: Paul Young et al.

;; TITLE OF INVENTION: PGRP-L Polynucleotides, Polypeptides, and Antibodies

;; FILE REFERENCE: PF513P1  
;; CURRENT APPLICATION NUMBER: US/10/068,956  
;; CURRENT FILING DATE: 2002-02-11  
;; PRIOR APPLICATION NUMBER: 60/149,715  
;; PRIOR FILING DATE: 1999-08-20  
;; PRIOR APPLICATION NUMBER: PCT/US00/22877  
;; PRIOR FILING DATE: 2000-08-18  
;; NUMBER OF SEQ ID NOS: 18  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 4  
;; LENGTH: 530  
;; TYPE: PRT  
;; ORGANISM: human  
;;  
US-10-068-956-4

Query Match 5.7%; Score 10; DB 15; Length 530;  
Best Local Similarity 100.0%; Pred. No. 0.68;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGHWWVGAHT 10  
Db 441 RGHWWVGAHT 450

RESULT 7

US-10-206-576-84  
; Sequence 84, Application US/10206576  
; Publication No. US20030017495A1

;; GENERAL INFORMATION:

;; APPLICANT: Choi et al.  
;; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
;; NUMBER OF SEQUENCES: 497  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: CD-R  
;; COMPUTER: Dell Latitude  
;; OPERATING SYSTEM: Windows 98  
;; SOFTWARE: ASCII Text

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/206,576  
;; FILING DATE: 29-Jul-2002  
;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 09/071,035  
;; FILING DATE: 1998-05-04  
;; APPLICATION NUMBER: US 60/046,655  
;; FILING DATE: 1997-05-16  
;; APPLICATION NUMBER: US 60/044,031  
;; FILING DATE: 1997-05-06  
;; APPLICATION NUMBER: US 60/066,009  
;; FILING DATE: 1997-11-14

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Hyman, Mark J.  
;; REGISTRATION NUMBER: 46,789  
;; REFERENCE/DOCKET NUMBER: PB369P1D1

;; INFORMATION FOR SEQ ID NO: 84:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 526 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear

;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 84:

US-10-206-576-84

Query Match

4.6%; Score 8; DB 12; Length 526;

Best Local Similarity 100.0%; Pred. No. 53;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ALPTEAAL 36  
Db 404 ALPTEAAL 411

RESULT 8

US-10-206-576-82  
; Sequence 82, Application US/10206576  
; Publication No. US20030017495A1

;; GENERAL INFORMATION:

;; APPLICANT: Choi et al.  
;; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
;; NUMBER OF SEQUENCES: 497  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: CD-R  
;; COMPUTER: Dell Latitude  
;; OPERATING SYSTEM: Windows 98  
;; SOFTWARE: ASCII Text

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/206,576  
;; FILING DATE: 29-Jul-2002  
;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 09/071,035  
;; FILING DATE: 1998-05-04  
;; APPLICATION NUMBER: US 60/046,655  
;; FILING DATE: 1997-05-16  
;; APPLICATION NUMBER: US 60/044,031  
;; FILING DATE: 1997-05-06  
;; APPLICATION NUMBER: US 60/066,009  
;; FILING DATE: 1997-11-14

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Hyman, Mark J.  
;; REGISTRATION NUMBER: 46,789  
;; REFERENCE/DOCKET NUMBER: PB369P1D1

;; INFORMATION FOR SEQ ID NO: 82:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 546 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear

;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 82:

US-10-206-576-82

Query Match 4.6%; Score 8; DB 12; Length 546;

Best Local Similarity 100.0%; Pred. No. 54;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ALPTEAAL 36  
Db 424 ALPTEAAL 431

RESULT 9

US-09-938-315-93  
; Sequence 93, Application US/09938315  
; Patent No. US20020091085A1

;; GENERAL INFORMATION:

;; APPLICANT: KAY, BRIAN K.  
;; SPARKS, ANDREW B.  
;; THORN, JUDITH M.  
;; QUILLIAM, LAWRENCE A.

Not part of seq 2  
at 30 of seq 2



DER, CHANNING J.  
TITLE OF INVENTION: SRC SH3 BINDING PEPTIDES AND METHODS OF  
ISOLATING AND USING SAME  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/938,315  
FILING DATE: 23-Aug-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Villacorta, Gilberto M.  
REGISTRATION NUMBER: 34,038  
REFERENCE/DOCKET NUMBER: 4980-007-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-09-938-315-93

Query Match 4.0%; Score 7; DB 9; Length 13;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 STRPLPP 112  
|||  
Db 2 STRPLPP 8

RESULT 10  
US-10-161-791-93  
; Sequence 93, Application US/10161791  
; Publication No. US20030186863A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/161,791  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,999  
FILING DATE: 16-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Misrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-202  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-10-161-791-93  
Query Match 4.0%; Score 7; DB 14; Length 13;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 106 STRPLPP 112  
|||  
Db 2 STRPLPP 8  
RESULT 11  
US-09-938-315-60  
; Sequence 60, Application US/09938315  
; Patent No. US20020091085A1  
; GENERAL INFORMATION:  
; APPLICANT: KAY, BRIAN K.  
; SPARKS, ANDREW B.  
; THORN, JUDITH M.  
; QUILLIAM, LAWRENCE A.  
; DER, CHANNING J.  
; TITLE OF INVENTION: SRC SH3 BINDING PEPTIDES AND METHODS OF  
; ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/938,315  
; FILING DATE: 23-Aug-2001  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Villacorta, Gilberto M.  
; REGISTRATION NUMBER: 34,038  
; REFERENCE/DOCKET NUMBER: 4980-007-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-938-315-60

```

Query Match	4.0%;	Score 7;	DB 9;	Length 31;
Best Local Similarity	100.0%;	Pred. No. 42;		
Matches	7;	Conservative 0;	Mismatches 0;	Indels 0;

Qy 106 STRPLPP 112  
db 16 STRPLPP 22

```

RESULT 12
US-10-161-791-60
; Sequence 60, Application US/10161791
; Publication No. US20030186863A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
;

```

Query Match 4.0%; Score 7; DB 14; Length 31;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels

Qy 106 STRPLPP 112  
Db 16 STRPLPP 22

```

RESULT 13
US-09-864-761-40010
; Sequence 40010, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40010
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006425.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EST HUMAN HIT: BF570453.1, EVALUATE 1.00e-11
; OTHER INFORMATION: SWISSPROT HIT: P52859, EVALUATE 4.00e-08
US-09-864-761-40010

```

Query Match 4.0%; Score 7; DB 9; Length 44;



Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 ALLGHRQ 64  
| | | | |  
Db 37 ALLGHRQ 43

RESULT 14  
US-10-029-386-30230  
; Sequence 30230, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 30230  
; LENGTH: 93  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR11.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2  
; OTHER INFORMATION: SWISSPROT HIT: P15442, EVALUE 1.60e+00  
US-10-029-386-30230

Query Match 4.0%; Score 7; DB 14; Length 93;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 LPTEAAL 36  
| | | | |  
Db 82 LPTEAAL 88

RESULT 15  
US-10-425-114-69630  
; Sequence 69630, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 69630  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73151C08\_FLI.pep  
US-10-425-114-69630

Query Match 4.0%; Score 7; DB 12; Length 103;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 TSSTRPL 110  
| | | | |  
Db 58 TSSTRPL 64

RESULT 16  
US-09-815-242-10432  
; Sequence 10432, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 10432  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-815-242-10432

Query Match 4.0%; Score 7; DB 9; Length 127;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VAIVGNY 26  
| | | | |  
Db 28 VAIVGNY 34

RESULT 17  
US-10-156-761-7920  
; Sequence 7920, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089

```
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7920
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7920

Query Match          4.0%; Score 7; DB 14; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      74 ALFDLLR 80
      |||||
Db      54 ALFDLLR 60

RESULT 18
US-10-424-599-192521
; Sequence 192521, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 192521
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(135)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_15869C.1.pep
US-10-424-599-192521

Query Match          4.0%; Score 7; DB 12; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 VGAHTLG 12
      |||||
Db      3 VGAHTLG 9

RESULT 19
US-10-424-599-190453
; Sequence 190453, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 190453
; LENGTH: 167
; TYPE: PRT
```

```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_14299C.1.pep
US-10-424-599-190453

Query Match          4.0%; Score 7; DB 12; Length 167;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34 AALRTVR 40
      |||||
Db      63 AALRTVR 69

RESULT 20
US-10-424-599-221448
; Sequence 221448, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 221448
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41998C.1.pep
US-10-424-599-221448

Query Match          4.0%; Score 7; DB 12; Length 182;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 LPTEAAL 36
      |||||
Db      143 LPTEAAL 149

RESULT 21
US-10-425-114-43280
; Sequence 43280, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43280
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700073330_FLI.pep
US-10-425-114-43280

Query Match          4.0%; Score 7; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
```



```
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      119 RTASARP 125
      |||||
Db      20 RTASARP 26

RESULT 22
US-10-282-122A-49294
; Sequence 49294, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49294
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-282-122A-49294

Query Match      4.0%;  Score 7;  DB 12;  Length 233;
Best Local Similarity 100.0%;  Pred. No. 2.3e+02;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      50 AGLLRPD 56
      |||||
Db      49 AGLLRPD 55

RESULT 23
US-10-425-114-48990
; Sequence 48990, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 48990
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700553168_FLI.pep
US-10-425-114-48990

Query Match      4.0%;  Score 7;  DB 12;  Length 233;
Best Local Similarity 100.0%;  Pred. No. 2.3e+02;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      118 ARTASAR 124
      |||||
Db      195 ARTASAR 201

RESULT 24
US-10-103-313-562
; Sequence 562, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ07C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 562
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-562

Query Match      4.0%;  Score 7;  DB 14;  Length 233;
Best Local Similarity 100.0%;  Pred. No. 2.3e+02;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      49 RAGLLRP 55
      |||||
Db      4 RAGLLRP 10

RESULT 25
US-10-369-493-20167
; Sequence 20167, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
```

;  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 20167  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: No. US20030233675A1toc punctiforme  
US-10-369-493-20167

Query Match 4.0%; Score 7; DB 15; Length 236;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 AGLLRPD 56  
| | | | |  
Db 42 AGLLRPD 48

RESULT 26  
US-10-424-599-231037  
; Sequence 231037, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 231037  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(314)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_50648C.1.pep  
US-10-424-599-231037

Query Match 4.0%; Score 7; DB 12; Length 314;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 VGAHTLG 12  
| | | | |  
Db 205 VGAHTLG 211

RESULT 27  
US-10-412-699B-1020  
; Sequence 1020, Application US/10412699B  
; Publication No. US20040045049A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: Zhang, James  
; APPLICANT: Fromm, Michael E.  
; APPLICANT: Heard, Jacqueline E.  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Adam, Luc J.  
; APPLICANT: Broun, Pierre E.  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddle, James S.  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Samaha, Raymond R.

;  
; APPLICANT: Pilgrim, Marsha L.  
; APPLICANT: Creelman, Robert A.  
; APPLICANT: DuBell, Arnold N.  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Kumimoto, Roderick  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants  
; FILE REFERENCE: MBI-0048CIP  
; CURRENT APPLICATION NUMBER: US/10/412,699B  
; CURRENT FILING DATE: 2003-04-10  
; PRIOR APPLICATION NUMBER: 09/394,519  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: 09/489,376  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: 09/506,720  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 09/533,030  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/533,392  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/533,029  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/532,591  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/533,648  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/713,994  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: 09/819,142  
; PRIOR FILING DATE: 2001-03-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2011  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1020  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-412-699B-1020

Query Match 4.0%; Score 7; DB 12; Length 361;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 AASAPQ 164  
| | | | |  
Db 257 AASAPQ 263

RESULT 28  
US-10-374-780A-525  
; Sequence 525, Application US/10374780A  
; Publication No. US20040019927A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, Bradley K  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Heard, Jacqueline E  
; APPLICANT: Haake, Volker  
; APPLICANT: Creelman, Robert A  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Adam, Luc J  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddle, James  
; APPLICANT: Broun, Pierre E  
; APPLICANT: Pilgrim, Marsha L.  
; APPLICANT: DuBell III, Arnold T  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
; FILE REFERENCE: MBI-0047 CIP  
; CURRENT APPLICATION NUMBER: US/10/374,780A  
; CURRENT FILING DATE: 2003-02-25  
; PRIOR APPLICATION NUMBER: 09/837,944



```
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 525
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Orthologous to G188
US-10-374-780A-525
```

```
Query Match          4.0%; Score 7; DB 15; Length 361;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      158 AASAOQP 164
        |||||
Db       257 AASAOQP 263
```

RESULT 29

```
US-10-210-130-118
; Sequence 118, Application US/10210130
; Publication No. US20040014053A1
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
```

```
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 118
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-130-118
```

```
Query Match          4.0%; Score 7; DB 15; Length 377;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      162 QPQTQPA 168
        |||||
Db       366 QPQTQPA 372
```

RESULT 30

```
US-10-287-274-393
; Sequence 393, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERETO
; FILE REFERENCE: ELITRA.008DV1
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 393
```

```
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-393

Query Match
Best Local Similarity 100.0%; Score 7; DB 14; Length 385;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 DALFDLL 79
|
Db 369 DALFDLL 375

RESULT 31
US-10-310-002-13
; Sequence 13, Application US/10310002
; Publication No. US20030125296A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Insulin-Responsive Sequence DNA Binding Protein-1 and Methods to
; TITLE OF INVENTION: Regulate Insulin-Responsive Genes
; FILE REFERENCE: E056 1010
; CURRENT APPLICATION NUMBER: US/10/310,002
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-310-002-13

Query Match
Best Local Similarity 100.0%; Score 7; DB 14; Length 387;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 TRPLPPA 113
|
Db 165 TRPLPPA 171

RESULT 32
US-10-424-599-144309
; Sequence 144309, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144309
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101324C.1.pep
US-10-424-599-144309

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 428;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 DALFDLL 79
|
Db 337 DALFDLL 343
```

```
RESULT 33
US-10-220-380-4
; Sequence 4, Application US/10220380
; Publication No. US20030113846A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; APPLICANT: NGUYEN, Dannie L.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: WALIA, Narinder K.
; APPLICANT: HAFALIA, April
; APPLICANT: TRIBOULEY, Catherine M.
; TITLE OF INVENTION: LIPID METABOLISM ENZYMES
; FILE REFERENCE: PI-0050 PCT
; CURRENT APPLICATION NUMBER: US/10/220,380
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 60/186,480; 60/190,415; 60/198,437
; PRIOR FILING DATE: 2000-03-02; 2000-03-17; 2000-04-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030113846A1 2690842CD1
US-10-220-380-4

Query Match
Best Local Similarity 100.0%; Score 7; DB 14; Length 432;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 ALPTEAA 35
|
Db 20 ALPTEAA 26

RESULT 34
US-10-425-114-43085
; Sequence 43085, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43085
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700685655_FLI.pep
US-10-425-114-43085

Query Match
Best Local Similarity 100.0%; Score 7; DB 12; Length 440;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      73 DALFDLL 79
      |||||
Db      349 DALFDLL 355

RESULT 35
US-10-425-114-59071
; Sequence 59071, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59071
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700042149_FLI.pep
US-10-425-114-59071

Query Match      4.0%;  Score 7;  DB 12;  Length 452;
Best Local Similarity 100.0%;  Pred. No. 4.1e+02;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      46 CAVRAGL 52
      |||||
Db      39 CAVRAGL 45

RESULT 36
US-10-424-599-234625
; Sequence 234625, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234625
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(495)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53895C.1.pep
US-10-424-599-234625

Query Match      4.0%;  Score 7;  DB 12;  Length 495;
Best Local Similarity 100.0%;  Pred. No. 4.4e+02;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      55 PDYALLG 61
```

```
Db      364 PDYALLG 370
      |||||

RESULT 37
US-10-425-114-59944
; Sequence 59944, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59944
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LJB3689-233-B2_FLI.pep
US-10-425-114-59944

Query Match      4.0%;  Score 7;  DB 12;  Length 500;
Best Local Similarity 100.0%;  Pred. No. 4.5e+02;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      105 SSTRPLP 111
      |||||
Db      16 SSTRPLP 22

RESULT 38
US-10-308-128-6
; Sequence 6, Application US/10308128
; Publication No. US20040033506A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL HUMAN MITOCHONDRIAL AND MICROSOMAL
; FILE REFERENCE: D0199 NP
; CURRENT APPLICATION NUMBER: US/10/308,128
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: U.S. 60/334,904
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-308-128-6

Query Match      4.0%;  Score 7;  DB 12;  Length 502;
Best Local Similarity 100.0%;  Pred. No. 4.5e+02;
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      29 ALPTEAA 35
      |||||
Db      90 ALPTEAA 96

RESULT 39
US-10-266-829-112
; Sequence 112, Application US/10266829
; Publication No. US20030220489A1
; GENERAL INFORMATION:
```

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 29 Human secreted proteins  
; FILE REFERENCE: P2041P1  
; CURRENT APPLICATION NUMBER: US/10/266,829  
; CURRENT FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 09/756,168  
; PRIOR FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: PCT/US00/19735  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 60/145,220  
; PRIOR FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 146  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 112  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-266-829-112

Query Match 4.0%; Score 7; DB 15; Length 505;  
Best Local Similarity 100.0%; Pred.No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 TARPSV 102  
|||||||  
Db 400 TARPSV 406

RESULT 40  
US-09-833-245-321  
; Sequence 321, Application US/09833245  
; Publication No. US20040010134A1  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 321  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-833-245-321

Query Match 4.0%; Score 7; DB 11; Length 509;  
Best Local Similarity 100.0%; Pred.No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 TARPSV 102  
|||||||  
Db 404 TARPSV 410

Search completed: May 18, 2004, 16:27:25  
Job time : 59 secs